The dynamic coupling interface of G-protein coupled receptors – A molecular dynamics simulations study.

D i s s e r t a t i o n

zur Erlangung des akademischen Grades

d o c t o r r e r u m n a t u r a l i u m

(Dr. rer. nat.)

im Fach Biophysik

eingereicht an der

Lebenswissenschaftlichen Fakultät

der Humboldt-Universität zu Berlin

von

Alexander Sebastian Rose, M.Sc.

Präsident der Humboldt-Universität zu Berlin

Prof. Dr. Jan-Hendrik Olbertz

Dekanin/Dekan der Lebenswissenschaftlichen Fakultät

Prof. Dr. Richard Lucius

Gutachter/innen:
1. Herr Prof. Dr. Klaus-Peter Hofmann
2. Herr PD Dr. Peter Hildebrand
3. Herr Prof. Dr. Peter Hegemann

Tag der mündlichen Prüfung: 12.05.2015

THE DYNAMIC COUPLING INTERFACE OF G-PROTEIN COUPLED RECEPTORS

A molecular dynamics simulations study

ALEXANDER S. ROSE

Lebenswissenschaftliche Fakultät
Humboldt-Universität zu Berlin

Winter 2014

ABSTRACT

To communicate with their environment, living cells feature receptors that provide a bridge across the enclosing membrane. The prevalent G protein-coupled receptors (GPCR) receive outside information through the binding of a ligand, which activates the receptor. During activation, an open intracellular crevice forms, to which a G protein ($G\alpha\beta\gamma$, G) can couple with its $G\alpha$ C-terminus. Binding to GPCRs triggers GDP/GTP exchange in the $G\alpha$ subunit of $G\alpha\beta\gamma$, necessary for further signal transfer within the cell. The coupling between receptor and $G\alpha\beta\gamma$ involves a series of dynamic structural changes that govern speed and specificity of the interaction. Here we used molecular dynamics (MD) simulations to elucidate molecular details of the GPCR $G\alpha\beta\gamma$ coupling process before and during GPCR $G\alpha\beta\gamma$ complex formation up to the GDP/GTP exchange.

The photoreceptor rhodopsin is tuned to provide fast and precise signal transfer. We investigated the role of its third intracellular loop in enabling signaling specificity and speed. Employing MD simulations, we find that the third intracellular loop is an intrinsically unstructured region, that — like the $G\alpha$ C-terminus — adopts a structured conformation upon complex formation. These findings corroborate a Fourier transform infrared spectroscopy study and allow structural interpretation of the spectroscopic data. Based on these data, we propose a stepwise mutual structuring scheme, that explains fast signal transfer by an increased likelihood of complex formation. The high signal fidelity is realized by the stepwise structuring allowing precise interrogation of a large interaction surface.

The β_2-adrenoceptor can couple to the G protein families G_s and G_i. However, agonists can bias the receptor to prefer coupling to one family. Here we utilized MD simulations and umbrella sampling calculations to determine structural differences in the binding modes of active β_2-adrenoceptor with G_s or G_i. Using peptides derived from the far C-terminus of the $G\alpha$-subunit, we find that G_i can stabilize active β_2-adrenoceptor in a conformation very similar to active rhodopsin binding to G_t, a close homologue of G_i. A comparison of the observed G_i binding mode with crystal structures of G_s in complex with active β_2-adrenoceptor suggests that the specific outward tilt of transmembrane helix 6 allows the receptor to distinguish between G_s and G_i. One hypothesis is that biased agonists control that tilt to realize their biasing ability.

Signal transfer from GPCR to $G\alpha\beta\gamma$ has to bridge the distance of 30 Å that separates the receptor binding crevice and the GDP/GTP binding pocket. The α5-helix in $G\alpha$ connects the binding pocket to

the receptor and its rotation has been directly linked to GDP release, but it is unclear how the rotation is induced. The intermediary complex between GPCR and G$\alpha\beta\gamma$ leading to GDP/GTP exchange still has GDP bound. For rhodopsin and the β_2-adrenoceptor, our MD simulations of the α5-helix show a rotation of the α5-helix resulting only from interactions with the receptor interface. We propose that the α5-helix rotates at the coupling interface of GPCR and G$\alpha\beta\gamma$ during the transition from the intermediary complex to the GDP-free complex of receptor and G$\alpha\beta\gamma$. Thereby the intermediate complex provides a structural scaffold in which the α5-helix can rotate, while the receptor and the G$\alpha\beta\gamma$ stay fixed, so that the rotation maximally effects the GDP binding pocket thus triggering GDP release.

ZUSAMMENFASSUNG

Um mit ihrer Umgebung zu kommunizieren verfügen lebende Zellen über Rezeptoren, welche die umschließende Membran überbrücken. Die vorherrschende G-Protein-gekoppelte Rezeptoren (GPCR) erhalten Informationen von Außerhalb durch Bindung eines Liganden, wodurch der Rezeptor aktiviert wird. Während der Aktivierung bildet sich innerzellulär ein offener Spalt, in den ein G-Protein ($G\alpha\beta\gamma$, G) mit seinem C-terminalen Ende koppeln kann. Die Bindung an einen GPCR führt in der $G\alpha$-Untereinheit vom $G\alpha\beta\gamma$ zu einen GDP/GTP-Austausch, welcher für die weitere Signalübertragung ins Zellinnere notwendig ist. Die Kopplung von Rezeptor und $G\alpha\beta\gamma$ umfasst eine Reihe von dynamischen strukturellen Änderungen, die Geschwindigkeit und Spezifität der Interaktion regeln. Hier haben wir MD-Simulationen (Molekulardynamik) verwendet, um die molekularen Details der GPCR $G\alpha\beta\gamma$ Kopplung vor und während der GPCR-$G\alpha\beta\gamma$-Komplexbildung bis hin zum GDP/GTP-Austausch zu untersuchen.

Der Photorezeptor Rhodopsin ist auf schnelle und präzise Signalübertragung abgestimmt. Wir haben untersucht, welche Rolle seine dritte intrazelluläre Schleife für das Zustandekommen dieser Eigenschaften spielt. In MD-Simulationen beobachten wir, dass die dritte intrazelluläre Schleife an sich unstrukturiert ist, aber — wie der $G\alpha$ C-Terminus — bei Komplexbildung eine strukturierte Konformation einnimmt. Diese Ergebnisse bestätigen eine FTIR-Studie (Fourier-Transformation-Infrarotspektroskopie) und ermöglichen die strukturelle Interpretation der spektroskopischen Daten. Basierend auf diesen Daten schlagen wir ein schrittweises, gegenseitiges Strukturierungsschema vor, welches eine schnelle Signalübertragung durch eine erhöhte Wahrscheinlichkeit der Komplexbildung erklärt. Die hohe Signaltreue wird durch die präzise Abfrage einer großen Interaktionsfläche während der stufenweisen Strukturierung realisiert.

Der β_2-Adrenozeptor kann an die G-Protein-Familien G_s und G_i koppeln. Allerdings gibt es steuernde Agonisten (sogenannte *biased agonists*), die den Rezeptor eine Familie bevorzugen lassen. Hier haben wir MD-Simulationen und Umbrella-Sampling-Berechnungen durchgeführt, um die strukturellen Unterschiede der Bindung aktiver β_2-Adrenozeptor mit G_s oder G_i zu bestimmen. Bei Verwendung von aus dem C-Terminus der $G\alpha$-Untereinheit abgeleiteten Peptiden fanden wir, dass G_i eine aktive Konformation des β_2-Adrenozeptors stabilisiert, die ähnlich zu aktivem, mit G_t gebundenem Rhodopsin ist, einem nahen Verwandten von G_i. Ausgehend vom Vergleich des beobachteten G_i-Bindungsmodus mit Kristallstrukturen von einem mit G_s komplexierten, aktiven β_2-Adrenozeptor schlagen wir vor, dass

die spezifische Auswärtsneigung von Transmembranhelix 6 es dem Rezeptor ermöglicht, zwischen G_s und G_i zu unterscheiden. Eine Hypothese ist, dass steuernde Agonisten diese Auswärtsneigungneigung nutzen, um ihre steuernde Fähigkeit zu realisieren.

Signalübertragung von GPCR zu $G\alpha\beta\gamma$ muss die Distanz von 30 Å überbrücken, welche den Bindungsspalt im Rezeptor und die GDP/GTP-Bindungstasche trennt. Die α5-Helix in $G\alpha$ verbindet die Bindungstasche mit dem Rezeptor und deren Drehung steht in direktem Zusammenhang mit der GDP-Freisetzung, aber es ist unklar wie die Drehung hervorgerufen wird. Den GDP/GTP-Austausch einleitend, hat der Zwischenkomplex von GPCR und $G\alpha\beta\gamma$ noch GDP gebunden. Für Rhodopsin und den β_2-Adrenozeptor zeigen unsere MD-Simulationen eine Drehung der α5-Helix als Resultat allein von Interaktionen mit der Kopplungsfläche des Rezeptors. Wir schlagen daher vor, dass die Drehung der α5-Helix in der Kopplungs-/Interaktionsfläche von GPCR und $G\alpha\beta\gamma$ während des Übergangs vom Zwischenkomplexes zum GDP-freien Rezeptor-$G\alpha\beta\gamma$-Komplex geschieht. Dadurch bietet der Zwischenkomplex ein Strukturgerüst, in welchem sich die α5-Helix drehen kann, während Rezeptor und $G\alpha\beta\gamma$ fixiert sind, sodass die Rotation maximale Wirkung auf die GDP-Bindungstasche hat und so die Freisetzung des GDP auszulösen vermag.

PUBLICATIONS

Some ideas and figures have appeared previously in the following publications or manuscripts that I have (co-)authored:

- MANUSCRIPT: **AS Rose**, U Zachariae, H Grubmüller, KP Hofmann, P Scheerer, PW Hildebrand. "The activating α5 helix switch is driven by the receptor G protein interface. (working title)"
- **AS Rose**, M Elgeti, U Zachariae, H Grubmüller, KP Hofmann, P Scheerer and PW Hildebrand. "Position of transmembrane helix 6 determines receptor g protein coupling specificity." *J. Am. Chem. Soc.* 136.32 (2014), pp. 11244–7. DOI:10.1021/ja5055109
- M Szczepek, F Beyrière, KP Hofmann, M Elgeti, R Kazmin, **AS Rose**, FJ Bartl, D von Stetten, M Heck, ME Sommer, PW Hildebrand, P Scheerer. "Crystal structure of a common GPCR-binding interface for G protein and arrestin." *Nat. Commun.*, 5.May (2014) p. 4801. DOI:10.1038/ncomms5801
- **AS Rose**, D Theune, A Goede, PW Hildebrand. "MP:PD—a data base of internal packing densities, internal packing defects and internal waters of helical membrane proteins." *Nucleic Acids Res.*, 42.Database issue (2014), pp. D347–51. DOI:10.1093/nar/gkt1062
- M Elgeti, **AS Rose**, FJ Bartl, PW Hildebrand, KP Hofmann, M Heck. "Precision vs flexibility in GPCR signaling." *J. Am. Chem. Soc.*, 135.33 (2013), pp. 12305–12. DOI:10.1021/ja405133k

Other publications that I have (co-)authored:

- J Ismer[1], **AS Rose[1]**, JKS Tiemann, A Goede, K Rother, PW Hildebrand. "Voronoia4RNA—a database of atomic packing densities of RNA structures and their complexes." *Nucleic Acids Res.*, 41.Database issue (2013), pp. D280–4. DOI:10.1093/nar/gks1061

1 equal contributions

- **AS Rose**, A Goede, PW Hildebrand. "MPlot—a server to analyze and visualize tertiary structure contacts and geometrical features of helical membrane proteins." *Nucleic Acids Res.*, 38.Web Server issue (2010), pp. W602–8. DOI:10.1093/nar/gkq401

- **AS Rose**, S Lorenzen, A Goede, B Gruening, PW Hildebrand. "RHYTHM—a server to predict the orientation of transmembrane helices in channels and membrane-coils." *Nucleic Acids Res.*, 37.Web Server issue (2009), pp. W575–80. DOI:10.1093/nar/gkp418

ACKNOWLEDGMENTS

First, I want to thank *Peter Hildebrand*, who basically introduced me to science and provided ongoing supervision and support. He trusted me with challenging projects, giving me the opportunity to grew both scientifically and personally. Without his group, formed during the years, I would have missed a lot of inspiring discussions and interesting side-projects. Thank you: *Jochen Ismer, Heiko Bittner, Johanna Tiemann, Dominik Theune* and our long term collaborator *Andrean Goede*.

Scientific works are normally products by many people, combining their knowledge to find things out, preferably new things. Honoring the collaborative effort that made this work possible, I want to thank, apart from Peter Hildebrand, my colleagues and close collaborators *Klaus-Peter Hofmann*, *Matthias Elgeti*, *Patrick Scheerer*. They generously shared their expertise on rhodopsin, G protein coupled receptors and science in general.

Many thanks also to my collaborators from Göttingen, *Ulrich Zachariae* and *Helmut Grubmüller*, who introduced me to the capabilities of molecular dynamics simulations and warned me about its limitations.

The *Institut für medizinische Physik und Biophysik* at the Charité—Universitätsmedizin Berlin was (and is) a great place to do research. While not a huge facility, there are many skilled and kind people: researchers, technicians and other staff alike. Their diverse research focuses and method interests allowed me to be in fruitful cooperations, which I hope will continue to flourish. I especially like to thank, if not already: *Franz Bartl, Florent Beyrière, Martin Heck, Roman Kazmin, Patrick Piwowarski, Martha Sommer, Christian Spahn, Michal Szczepek*.

An honorable mention goes to all the *people* who contributed to the many open source software tools that were essential to perform my research. Thank you for sharing the code, I will try to pay you back in turn.

Funding for this work was provided by the Deutsche Forschungsgemeinschaft and the European Research Council. The computer time necessary for this project was provided in part by the "Norddeutscher Verbund fur Hoch- und Höchstleistungsrechner".

And last but not least, I would like to thank my *parents* and *siblings* for supporting me in so many way. They provided much welcomed distractions from research and patiently endured my ignorance for anything not thesis related during the final months.

CONTENTS

LIST OF SCHEMES

LIST OF EQUATIONS

LIST OF TABLES

CODE LISTINGS

1 INTRODUCTION

1.1 BIOLOGICAL BACKGROUND

1.1.1 *Overview of signal transduction by GPCRs*

Living cells must communicate with their environment to receive sensory stimuli and to cooperate with other cells, especially in multicellular organisms. To perform these functions, cells utilize proteins that are embedded in the cell membrane thus bridging the inside of the cell (or cellular compartment) with the outside. A widely deployed class of proteins are G protein-coupled receptors (GPCRs) stretching with seven transmembrane helices (7-TM) across the lipid bilayer that surrounds the cell [1, 2, 3]. One side of these receptors sticks out into the extracellular medium, while the other is in contact with the cell interior. Given this topology, GPCRs can transmit signals into the cell by responding to extracellular signals. This is for most systems a small ligand molecule that approaches the receptor (R) from the outside and by binding activates the receptor. The so activated receptor changes at its intracellular side, allowing receptor binding proteins like G proteins or arrestins to couple to the receptor and relay the signal to effectors, eventually leading to a cellular response (Figure 1). For example, the β_2-adrenoceptor (β_2AR), responsible for smooth muscle relaxation by adrenaline, is activated by diffusible ligands and couples to two different heterotrimeric G proteins (G$\alpha\beta\gamma$, families G_i, G_s, G_q, $G_{12/13}$): G_s and G_i [4]. Rhodopsin from the visual system on the other hand features a covalently bound ligand that activates upon light absorption [1].

Signal transduction. Generally, different GPCRs can receive a large variety of extracellular signals in the form of various ligands and transduce them into the cell, for example to G proteins. Ligands can be distinguished by their effect on receptor activation and subsequent signal transduction through G protein coupling. Agonists have an activating effect, whereas inverse agonists stabilize the inactive receptor. The agonist-induced active state of GPCRs can couple to G proteins and catalyze the nucleotide exchange of GDP for GTP in the α-subunit of the heterotrimeric G protein [12]. After GTP-binding the G$\alpha\beta\gamma$ holoprotein dissociates from the receptor. Eventually, the Gα-subunit and the G$\beta\gamma$-complex separate and trigger various signaling pathways by coupling to a downstream intracellular effector (Figure 1). More specifically, a single receptor type can activate distinct intracellular downstream signaling proteins, includ-

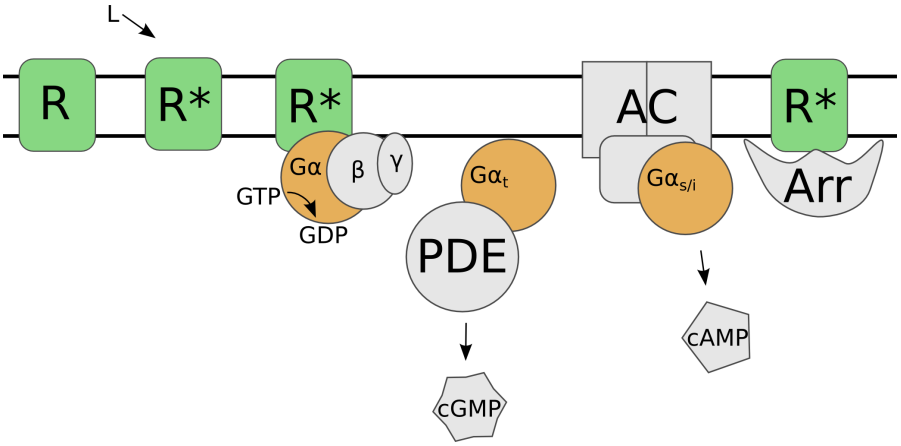

Figure 1: GPCR signaling pathways. Signaling starts with a receptor (green, R) being activated (R*) by a ligand (L). A heterotrimeric G protein (orange, $G\alpha\beta\gamma$) may then be activated by GDP to GTP nucleotide exchange upon binding to R* [5]. Eventually the G protein dissociates and in its GTP bound state, the α-subunit further transmits the signal, depending on its type [6]. Transducin ($G\alpha_t$) activates cGMP phosphdiesterase (PDE), resulting in rising cyclic guanosine monophosphate (cGMP) levels [7]. Membrane bound adenylate cyclase sub types (AC) can be activated by $G\alpha_s$ and inhibited by $G\alpha_i$, resulting in respective changes of cyclic adenosine monophosphate (cAMP) levels [8]. Arrestin (Arr) generally shuts down the G protein dependent pathway and targets receptors for internalization but can also trigger its own signaling pathway [9, 10]. The remaining G_q pathway via phospholipase C (PLC) is not shown [11].

ing the heterotrimeric G proteins and arrestins (arrestin 1–4). This binding promiscuity results in different cellular and physiological responses [13].

Pharmacological side effects. The central role of GPCR in signal transduction across the cell membrane makes them a prime target for pharmaceuticals [14, 15]. GPCRs are one of the main mechanisms by which cells are regulated in a multicellular organism. Pharmaceutical molecules can act similar to endogenous ligands, with the aim to positively influence cells exhibiting pathological behavior. A main challenge in the approach are the multifaceted properties of GPCRs. The receptors bind to different ligands, whose effects fill the whole spectrum from complete deactivation to full or even over-activation of the receptor [16, 17, 18, 19]. Besides ligands, other factors that influence receptor activation and G protein coupling include receptor dimerization [20], cellular localization [21], sodium ions [22], pH value [23] or lipid composition [24]. Therefore, influenced by the environment, the complex relationship between ligand (pharmaceutical) and cellular effect (i. e. between stimulus and response) is likely a source for (mostly unintended) side effects of pharmaceuticals.

Coupling promiscuity mechanism. Elucidating the structural details of receptor G protein coupling may help to entangle some of the

intricacies governing stimulus and response in GPCR signaling. Especially the molecular mechanism that underlies the coupling promiscuity observed for various receptors is a current focus of receptor research. The ability of ligands to bias either towards arrestin- or to G protein-specific signaling has been linked to different conformations observed for the active receptor (R*) [25, 26]. Specifically, it has been suggested that there exist different R* conformations that exhibit distinct G protein and arrestin binding properties. For the human β_2-adrenoceptor some structural details of how different ligands can bias β_2AR toward the arrestin or G protein signaling pathway have recently been elucidated [27]. Even less is known about the structural mechanism utilized by receptors to distinguish between different G proteins for coupling, by which the receptor can selectively activate distinct downstream effector pathways. For example, β_2AR can activate both the stimulatory G_s or the inhibitory G_i protein signaling pathway [4], but the structural details by which the receptor achieves this are unknown.

1.1.2 *Activation of GPCRs*

Before G protein or arrestin can bind to a GPCR, the receptor needs to be activated, a process involving major structural changes. A hallmark of receptor activation is the outward tilt of transmembrane helix 6 (TM6). First identified by pioneering EPR work on the retinal photoreceptor rhodopsin [28, 29], the TM6 movement has been confirmed by an array of R* structures, obtained by X-ray crystallography for rhodopsin [30, 31], the β_2-adrenoceptor [32] and the muscarinic M2 receptor [33]. These R* structures have undergone significant structural changes compared to the inactive R structures and bind agonistic ligands, G protein, fragments of the Gα-subunit or nanobodies (see ref [34]). In addition to the TM6 movement, changes to TM5, comprising helix elongation and rigid body movements, are observed in the crystal structures of R* in complex with Gα or parts thereof [31, 32]. Together, these major structural rearrangements result in the formation of a open intracellular binding crevice. The helix movements involve a number of microswitches of highly conserved residues and motifs [1]. One of them is the arginine from the conserved E(D)R$^{3.50}$Y motif[1], which is buried in the inactive receptor state but gets exposed in the active state where it forms the floor of the created binding crevice (Figure 2). Stabilization of the inactive receptor is facilitated by the hydrogen bond network of deprotonated E$^{3.49}$, R$^{3.50}$ and E$^{6.30}$ tethering TM3 and TM6 together [36, 37, 38]. Upon receptor activa-

1 R135 in rhodopsin or generally R$^{3.50}$, utilizing the Ballesteros-Weinstein scheme where the most conserved residue in each TM among GPCRs is designated x.50, where x is the TM number and .50 is used for reference. Other residues on the same TM are numbered relative to the x.50 residue [35].

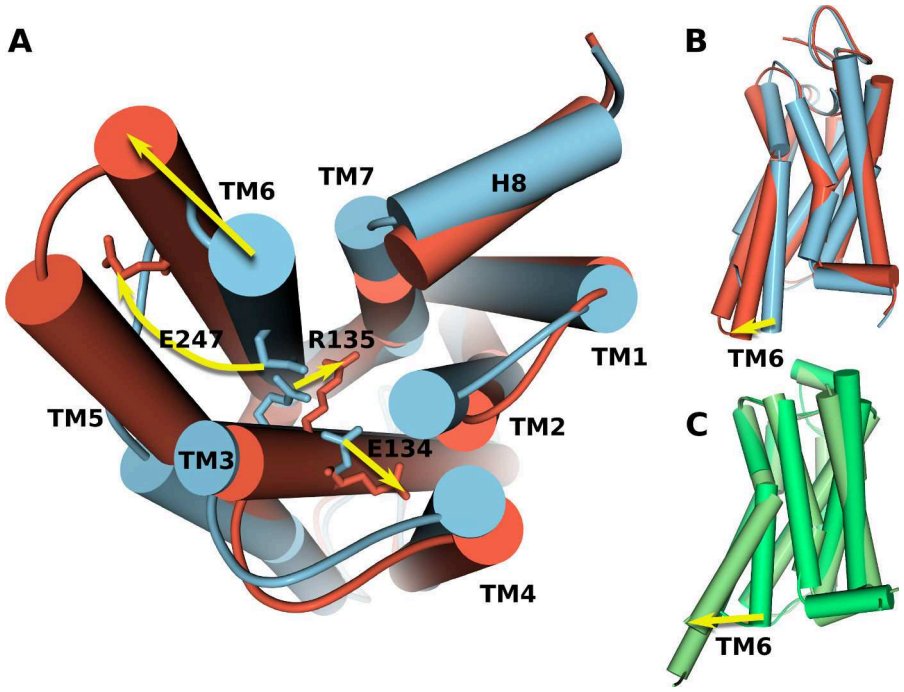

Figure 2: GPCR activation requires major structural changes and a series of microswitches involving conserved residues. (A) Comparison of inactive rhodopsin (blue) and active rhodopsin (red). Upon activation, TM6 tilts outwards and the hydrogen bond network involving E134, R135 and E247 is broken, resulting in an open intracellular crevice at the receptor surface. Side-view showing the major structural changes in the helical bundle of rhodopsin (B) and β_2-adrenoceptor (C) crystal structures upon activation. Note the large outward tilt of TM6 by a rigid body movement of the helix (yellow arrow), distinguishing the inactive and active receptor states. The crystal structures shown are RhR [39] (blue), RhR* [30](red), β_2AR [40] (dark green) and β_2AR* [32] (light green).

tion, proton uptake by $E^{3.49}$ disrupts this network allowing the new $R^{3.50}$ conformation.

Rhodopsin, while an archetype of class A GPCRs, is also a special case as it features a covalently bound light-sensitive ligand retinal, attached via a Schiff base linkage. Retinal tightly controls the transition/switch from inactive (RhR) to active (RhR*) rhodopsin. In contrast to the diffusible ligands employed by other GPCRs, retinal serves a dual role. In its *11-cis* form it acts as a strong inverse agonist but the absorption of a photon induces the isomerization of retinal into its *all-trans* form, which is a potent agonist for rhodopsin. This isomerization puts structural stress on the retinal binding pocket to which the receptor reacts by relaxing into (photo)intermediate states (batho-, lumirhodopsin), which eventually lead to receptor activation [41, 42]. By light activation, the receptor leaves its inactive dark state and enters an equilibrium of so-called metarhodopsin states (MI, MIIa, MIIb, $MIIbH^+$, Scheme 1), of which the last ($MIIbH^+$) can couple productively to G. Formation of the active $MIIbH^+$ state is ac-

companied by the protonation of D134 ($D^{3.49}$) [36]. Already during the transition to MIIb, the hydrogen bond network between R135 and D247 is disturbed, severing the ties between TM3 (R135) and TM6 (D247), a prerequisite for TM6 outward movement [43]. Once activated, rhodopsin can in turn activate up to 10^3 G proteins per second and more under optimal conditions [44]. Eventually, the Schiff base linking retinal to rhodopsin is hydrolyzed and *all-trans* retinal leaves through a hydrophobic channel [45, 46]. This newly formed apoprotein opsin exists in an equilibrium of active (Ops*) and inactive (Ops) states.

$$\mathrm{RhR} \xrightarrow{\lightning\nu} \mathrm{batho} \rightarrow \mathrm{lumi} \rightarrow \mathrm{MI} \rightleftharpoons \mathrm{MIIa} \rightleftharpoons \mathrm{MIIb} \rightleftharpoons \mathrm{MIIbH^+}$$

Scheme 1: Rhodopsin activation scheme. After RhR absorbs light ($\lightning\nu$), the photointermediates batho- and lumirhodopsin (batho, lumi) are transiently traversed to culminate in the coupled equilibria of metarhodopsin states (MI, MIIa, MIIb, $\mathrm{MIIbH^+}$)

In contrast to rhodopsin, most other GPCRs like the β_2-adrenoceptor show considerable constitutive activity [47] and the process that leads to receptor activation is not known in the detail to which the metarhodopsin states have been elucidated. However, experimental data suggests that non-rhodopsin GPCRs have structural equivalents to the metarhodopsin and opsin states [1, 48, 49]. Thus, after the photoactivation of rhodopsin is completed, retinal acts similar to the more common diffusible ligands that bind to other GPCRs for example to the β_2-adrenoceptor. The crucial protonation event of $D^{3.49}$, known from rhodopsin, has also been shown for the β_2-adrenoceptor [50] as well as for the α_{1B}-adrenergic receptor [51]. This protonated species constitutes the active receptor: $\mathrm{MIIbH^+}$ in the context of metarhodopsin states but generally referred to as R* or specifically as RhR* and β_2AR* for active rhodopsin and β_2-adrenoceptor, respectively.

It is important to note that the active receptor R* at the end of the activation process by no means exhibits a globally fixed or rigid structure, as being an endpoint may suggest or as one might be inclined to deduce from the respective crystal structures. For example, loop motions generally can happen within nanoseconds and larger collective motions can already occur on the microsecond timescale [52]. So, while the rigid body movement of TM6 is a major distinguishing feature of the inactive and active receptor states, it does not imply that TM6 (or specifically its intracellular extension) remains at a single position in the active state. There are additional changes at the conserved $NP^{7.50}$xxY motif and TM5 observed during the activation process that are crucial for the conformational change of TM6 [1, 53, 54]. Thus the TM6 outward tilt is not the sole determinant for activation nor does TM6 exhibit a purely switch-like behavior. But TM6

seems to exist in a spectrum of outward tilts even after activation. X-ray crystallography and spectroscopic studies suggest that an intracellular binding partner is required to restrict the TM6 ensemble to a single conformation [31, 32, 33, 27, 55, 56].

1.1.3 *Coupling interface between GPCR and G protein*

Upon receptor activation, an open intracellular crevice is formed when TM6 tilts outward. This crevice is lined by the intracellular ends of the TM helices comprising the 7-TM bundle, their connecting intracellular loops (ICL1–3) and helix 8, which lies with its axis parallel on the lipid bilayer. The floor of the crevice is formed by the highly conserved $R^{3.50}$ from the E(D)$R^{3.50}$Y motif. Crystal structures of active rhodopsin in complex with a peptide derived from the C-terminus of the G protein α-subunit (RhR*•Gα_tCT) and of agonist-bound β_2-adrenoceptor in complex with the G$\alpha_s\beta\gamma$ holoprotein (β_2AR*•G$\alpha_s\beta\gamma$) reveal atomic details of the coupling interface between receptor and G protein [31, 32] (Figure 3). On the receptor side, the ICL3 loop connection TM5 with TM6 and the conserved $R^{3.50}$ are major determinants of G protein binding to the receptor. Additionally, contacts to the ICL2 and the NP$^{7.50}$xxY(x)$_{5,6}$F region are observed in the RhR*•Gα_tCT and β_2AR*•G$\alpha_s\beta\gamma$ crystal structures.

A major role in coupling of receptor and G protein is played by the third intracellular loop (ICL3), which is located adjacent to the conserved E(D)$R^{3.50}$Y motif and connects TM5 with TM6 [57, 58, 31, 32]. For most GPCRs, crystallographic studies indicate, that ICL3 is likely a long, unstructured protein segment: High temperature factors (or B factors) or a lack of electron density for atoms within ICL3 indicate a high flexibility of ICL3 and the in the intracellular ends TM5 and TM6. Moreover, many GPCRs could so far not be crystallized and resolved with their native ICL3 intact. To circumvent ICL3 flexibility, truncation of ICL3 or substitution by fusion proteins (like the T4-lysozyme) to minimize flexibility and to enhance crystallographic contacts have proven successful [59]. This picture changes for crystal structures that depict an active GPCR in complex with an intracellular binding partner like G proteins, mimicking proteins or fragments. There, no fusion proteins are attached to ICL3 as they would likely interfere with G protein binding. As a result, the intracellular ends of TM5 and TM6 exhibit helical extensions, leaving only outer ICL3 segment unresolved, for example in the β_2-adrenoceptor [60, 32] or in the M2 muscarinic receptor [33]. Such an ordering is supported by spectroscopic studies [27, 55, 56]. In the rhodopsin case ICL3 is comparatively short and stabilized by contacts in the crystal lattice, which leads to very similar structures of RhR* and RhR*•Gα_tCT, both exhibiting a well-ordered ICL3 [39, 30].

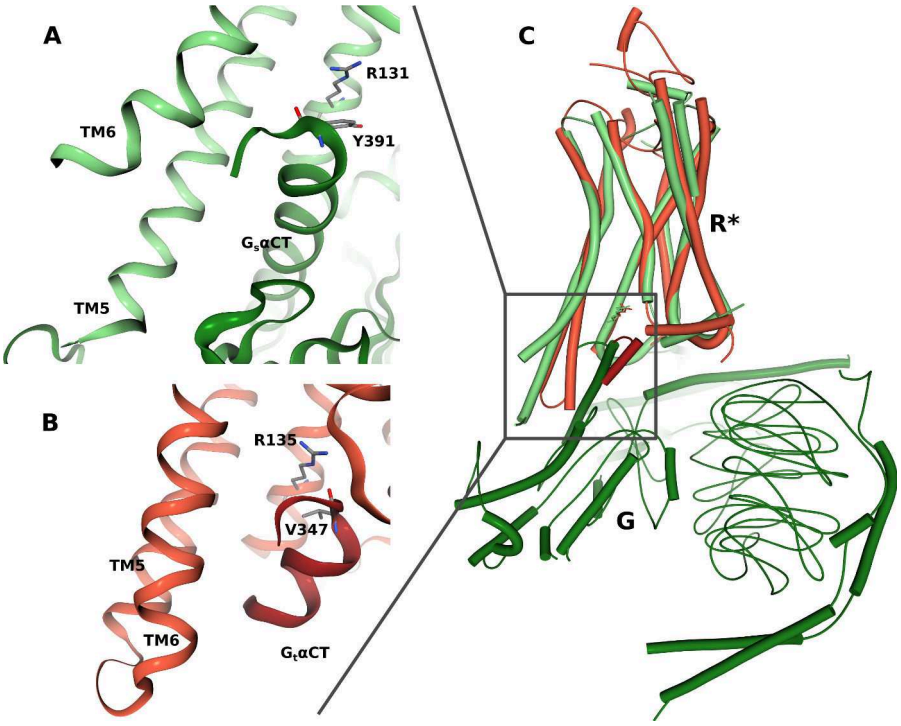

Figure 3: GPCR coupling interface between R* and GαCT as seen in the crystal structures of β_2AR*•G_s (green) and RhR*•$G\alpha_t$CT (red) complexes. (A) The Y391 at the tip of $G\alpha_s$CT forms a cation-π interaction with $R131^{3.50}$ of β_2AR*. (B) In RhR*•$G\alpha_t$CT, a hydrogen bond between the carbonyl oxygen of V347 at the tip of $G\alpha_t$CT and $R135^{3.50}$ is formed. Note how TM5/6 form a large contact area between R* and GαCT in both complexes, the hydrophobic patch. (C) Overview of the complexes showing the location (gray box) of the coupling interfaces which are magnified in (A) and (B).

The findings of crystallographic and spectroscopic studies on ICL3 and the adjacent TM6 make ICL3 a strong candidate for an intrinsically disordered protein region. Such regions are characterized by a lack of intrinsic order but upon binding to an interaction partner they may transition to an ordered state. The ICL3 shows this characteristic and is thus likely to represent an intrinsically disordered region within GPCRs [61, 62]. Such regions have been shown to be of high functional relevance for the binding and regulation of protein-protein interactions in many different biological systems [63]. However, for the coupling of receptor and G protein it remains to be elucidated to what extent the ICL3 flexibility and eventual ordering is exploited. The intrinsically disordered characteristic of ICL3 may play a role in binding as well as in signal transduction (i. e. G protein activation).

Gα C-terminus

The C-terminus of the G protein α-subunit (GαCT) has repeatedly been shown to be a main interaction site with the GPCR governing coupling fidelity and specificity [70, 71, 72, 73, 74]. Determina-

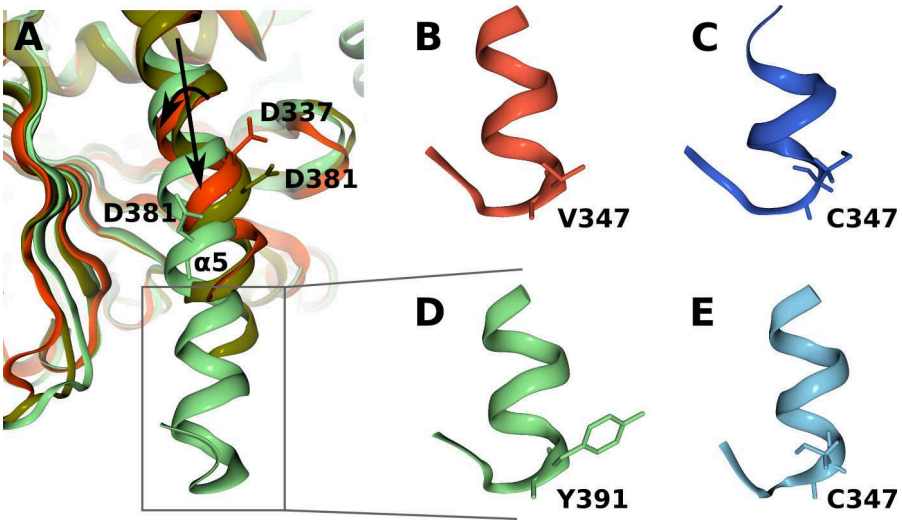

Figure 4: GαCT structure. (A) Crystal structures of a GDP-bound $G\alpha_{i/t}\beta\gamma$ chimera (orange) [64] or of GTPγS bound G_s (olive) [65] leave the GαCT unresolved (gray box). But in the β_2AR*•G_s complex (green) [32] $G\alpha_s$CT is well-ordered. Moreover, the α5-helix, which ends in the GαCT, is rotated and translated along its axis (black arrows) as indicated by the movement of D337/381 (Note that to adequately compare the α5-helices the Gα Ras domains where superposed on their β-sheets). Upon binding to R* GαCT is structured forming a continuous α-helix capped by a reverse turn motif as shown in numerous structures: (B) high affinity $G\alpha_t$CT from RhR*•$G\alpha_t$CT crystal structure (red) [31, 66, 67], (C) $G\alpha_t$CT NMR structure while bound to RhR* (blue) [68], (D) $G\alpha_s$CT from the β_2AR*•$G\alpha_s$CT crystal structure (green) [32], (E) another $G\alpha_t$CT NMR structure while bound to RhR* (light blue) [69].

tion of the crystal structures of the RhR*•$G\alpha_t$CT [31] and β_2AR*•$G\alpha_s\beta\gamma$ [32] complex then showed the coupling interface between receptor and GPCR in atomic detail. A key feature was thereby the specific interaction between the tip of GαCT and the conserved arginine $R^{3.50}$. Another articulate feature of the complexes is the hydrophobic patch between GαCT and ICL3, providing another hint at the importance of ICL3 (Figure 3). Additional contacts are observed from ICL2 to the αN-β1 and β2-β3 loops, both from the Gα Ras domain [32]. Thus GαCT interacts directly with the receptor binding crevice in a well-ordered conformation. But in its unbound state GαCT is unstructured. Only upon binding to a receptor, a continuous α-helix forms, which is capped by a reverse turn motif, and numerous NMR and X-ray structures of different Gα C-termini revealed a remarkable preservation of that capping motif (Figure 4) [68, 69, 31, 66, 67, 32]. Also, as shown for the rhodopsin system, peptides of this motif can stabilize the active receptor conformation, serving as a surrogate of complete G proteins [31, 66]. Comparing the C-termini of $G\alpha_s$ and $G\alpha_t$ shows that they share the same structure despite a different amino acid sequence, for example a tyrosine at the tip of $G\alpha_s$ instead

of a glycine. Moreover, for another GPCR binding partner, arrestin, the so-called "finger loop" is considered to directly interact with the receptor binding crevice. A recent crystallographic study shows a peptide derived from the arrestin "finger loop" (ArrFL) bound to the same intracellular receptor crevice employed by GαCT. They further suggest a fundamental connection between the sequences of GαCT and ArrFl segments: a common binding motif characterized by similar residues at key positions [75].

So far only crystal structures of RhR*•Gα_tCT and β_2AR*•G$\alpha_s\beta\gamma$ have been determined, raising as many questions as they answered. The β_2AR*•G_s complex structure provides a view at the nucleotide-free state of the GPCR G protein complex after GDP release and before GTP uptake and complex dissociation, two processes that still require structural elucidation. While β_2AR* is known to couple to G_s and G_i, no complex depicting the G_i binding mode is available. But there is the structure of active rhodopsin in complex with a G_t C-terminal peptide, which is a close homologue of G_i. Comparison of these complex structures shows that the Gα C-termini bind to the same binding crevice of the respective receptor. However, the TM6 tilt in the RhR*•Gα_tCT and β_2AR*•G$\alpha_s\beta\gamma$ structures differ considerably. It is not clear what the source of this differences is exactly as the structures show two different receptors with two different G proteins (or fragments thereof) obtained under unique crystallographic conditions. The distinct TM6 tilts may originate from the use of peptide versus holoprotein but they may also characterize two receptor conformations utilized to distinguish between G_s and G$\alpha_{i/t}$ which β_2AR* must accommodate both.

1.1.4 *Mechanistic aspects of signal transduction*

For signal transduction of GPCRs to G proteins, a signaling complex consisting of the receptor and the heterotrimeric G protein needs to be formed. The dynamic assembly and activation of such GPCR signaling complexes requires a complex series of conformational changes in both receptor and G protein [76, 77] (Figure 5). Generally, signal transduction from receptor to G protein is performed by the receptor triggering nucleotide exchange, GDP for GTP, in the G protein. Thereby GDP is released and GTP taken up. Thus activated, the G protein sends the signal further down the signaling cascade (Figure 1).

Upon formation of the nucleotide-free R*•G complex, a number of changes are known to occur within the G protein. First, this "Empty Pocket" state was found to have a conformationally dynamic Gα subunit [79]. Using site-directed spin labeling (SDSL) and electron paramagnetic resonance (EPR), the C-terminal α5-helix of Gα was found to undergo a rigid body movement in form of a rotation and a translation with respect to its helix axis (Figure 4A) [80]. A similar rigid

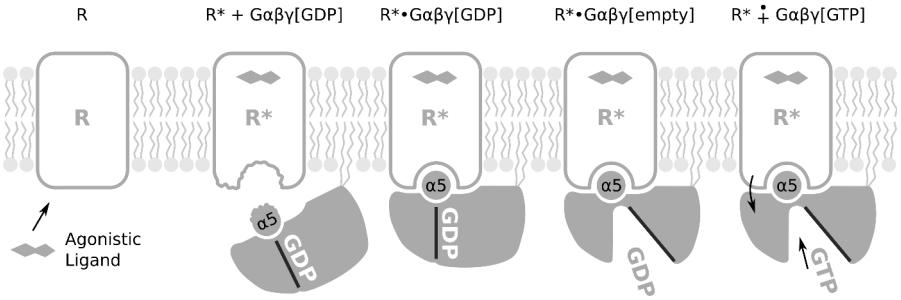

Figure 5: Sequence of complex conformations/states during GPCR G protein signaling, including intermediary states [76, 78]. An inactive receptor (R) is activated by an agonistic ligand, yielding an active receptor (R*), that presents a flexible binding crevice to a GDP-bound G protein (Gαβγ[GDP]) featuring an unstructured α5 C-terminus (α5). Upon formation of the R*• Gαβγ[GDP] intermediate and the following transition to the nucleotide-free R*•Gαβγ[empty] complex, the coupling interface becomes structured and α5 is displaced (see running text for details). During formation of R*• Gαβγ[empty], GDP is released and replaced by GTP, resulting in a dissociated Gαβγ[GTP] separated from R*.

body movement and, additionally, a mechanistic role of the GPCR coupling interface were suggested by a molecular docking and modeling study combining the Ops*•$G\alpha_t$CT [31] and $G\alpha_{i/t}\beta\gamma$ [64] crystal structures [78]. By applying SDSL and EPR, a large movement of the all-helical domain was observed opening an interdomain interface in the Gα subunit upon interaction with R* [81]. Similar movements are suggested by low resolution electron microscopy (EM) structures [82]. These earlier observations were confirmed by the X-ray structure of β_2AR* in complex with a nucleotide-free $G\alpha_s\beta\gamma$ [32] and accompanying hydrogen-deuterium exchange (HDX) measurements [83]. An increase in hydrogen deuterium exchange for nucleotide-free G_t in complex with RhR* has been suggested to allow GTP uptake and mediation of complex dissociation [84].

Molecular dynamics studies and *in silico* energetic analysis were used to shed light on the receptor G protein interaction for systems where other structural information is rare. The crystal structure of the β_2AR*•$G\alpha_s\beta\gamma$ complex provides a starting position for MD simulations of the coupling interface between receptor and G protein. Using homology modeling to create an initial model, coupling of G_i to the cannabinoid CB_1 [85] and CB_2 [86] receptor was investigated with MD simulations. The work with CB_1 used mutagenesis experiments and MD simulations to identify a number of residues in α5 contacting intra-cellular loop (ICL) 2 and 3, that reduce coupling when mutated to alanine. The latter study with CB_2 observes a reorientation of $G\alpha_i\beta\gamma$, more precisely a tilting of the α5-helix and an outward movement of trans-membrane helix (TM) 5, both of which are compatible with the cross-link sites found in the same work. *In silico* energetic analysis of a β_2RhR*•$G\alpha_i\beta\gamma$ complex based on the β_2AR*•

$G\alpha_s\beta\gamma$ crystal structure finds energetic (and quantitatively describes) changes upon α5 R* interaction [87].

α5-helix

The role of the α5-helix in G protein activation is only partially understood despite the determination of important conformational states, in particular the solvated GDP-bound and the nucleotide-free receptor-bound state. Extensive mutational studies on residues in the α5-helix have underlined the role of the α5-helix in allosterically connecting receptor and nucleotide. The T329A mutation in α5 of $G\alpha_i$ allosterically changes the switch I region, moving away from the nucleotide presumably destabilizing it [88]. Perturbation of the hydrophobic patch created by the conserved phenylalanine F336 in the α5-helix (e. g. F336Y mutant), suggests F336 acts as a relay between the α5-helix and the GDP binding region of heterotrimeric G_i protein α-subunit [89]. Also in $G\alpha_i$, a K345L mutation within α5 reduces nucleotide exchange rates, hinting at a transient interaction between the adjacent P-loop and the switch I region that contributes to the allosteric network between receptor and nucleotide binding pocket [90].

The importance of α5 in GDP release is undisputed, and rotation of α5 has been identified as a crucial structural change. Understanding the role of the active receptor in inducing the structural changes of α5 is complicated by the lack of structural data on intermediate R* G interactions. The existence of R*•G[GDP] intermediate was experimentally shown by a study on the kinetics of Gt activation by RhR* in disc membranes [44, 78]. Generally, the R*•G[GDP] intermediate's defining feature is that GDP is still bound in the nucleotide binding site of the $G\alpha$ subunit after complex formation with the active receptor. Recently, HDX experiments with the β_2AR* G_s system suggest that GDP-bound $G\alpha_s$ couples to β_2AR* mainly through $G\alpha_s$CT but weaker than in the nucleotide-free complex [83]. The intermediary R*•G[GDP] is the first stable/specific complex of receptor and G protein in the sequence of events during receptor triggered G protein activation (Figure 5). A possible mechanistic role for the intermediary R*•G[GDP] is that α5-helix rotation starts with the intermediary R*•G[GDP] complex but not before, thus separating initial R* and G complex formation and nucleotide release [78]. The dynamic and structural details of such a transition are, however, not known.

1.1.5 *Peptides as G protein surrogates*

Cellular processes often rely on small peptides, including in gene regulation [91], immune response [92], cell signaling or signal transduction [93, 94]. Following their endogenous models, synthetic peptides, their derivatives and peptide mimetics are used for example as antibiotics [95]. However, they have also been found a useful tool

for studying protein-protein interactions by modulating or inhibiting them [96, 97, 98]. Especially research on signal transduction through GPCRs has profited from the use of synthetic peptides. The signal transmission from activated receptors to their signaling proteins has been studied by synthetic peptides derived from GPCRs [99], G proteins ($G\alpha\beta\gamma$, G) [100, 74, 68, 69, 101, 31] and arrestin [102, 75]. These peptides were first derived from regions all over the holoprotein to identify regions crucial for complex formation. In case of rhodopsin, $G\alpha_tCT$ stabilizes the same active $MIIbH^+$ receptor conformation as G_t [100]. Intrinsically unstructured regions were identified as crucial, because they change their conformation upon complex formation (i. e. become structured) [103]. The so identified peptides therefore mimic the binding of an interaction partner and are as such used to investigate dynamic and mechanistic aspects of signal transmission that rely on complex formation.

In the context of molecular dynamics simulations, utilizing peptides opens up new opportunities when they are used to replace a (large) holoprotein. Biologically relevant processes involving conformational changes of the protein structure mostly occur on timescales ranging from microseconds to minutes, that only begin to become accessible to molecular dynamics simulations. By using a peptide, any constraints depending on the larger holoprotein from which the peptide is derived become obsolete. In a complex of an active receptor and a $G\alpha$CT peptide this lowers the number of interactions in the complex and thus also the interaction energy allowing easier transitions between different states of the complex. Additionally, a derived peptide is automatically much smaller, resulting in a simulation system with less atoms. So, using peptides as surrogates in MD simulations potentially requires smaller timescales due to lower transition barriers and allows longer simulations as less atoms need to be simulated.

1.2 MD SIMULATIONS

While static structures are determined for more and more proteins, in the end, conformational dynamics determine the protein's functions. MD simulations provide dynamic information at angstrom and picosecond resolution at the same time, which is outside the scope of any experimental method. Just now, with continuing advances in computational and experimental methods, it becomes possible to directly compare computational results to experiments [52, 104]. For example, a combination of temperature jump and fluorescence experiments can detect conversion rates and the existence of intermediate states of fast folding proteins in the microsecond timescale [105]. These observations can then be compared with data from MD sim-

ulations of fast-folding proteins, simulated for several hundred microseconds [106].

Since the first biologically relevant simulations of the bovine pancreatic trypsin inhibitor (PTI) [107], the field has made remarkable progress. It is now possible to simulate large biomolecules in explicit solvents and lipids, e. g. membrane proteins [108], do computational electrophysiology [109] or perform folding simulations for milliseconds [110].

1.2.1 *Augmenting static structures*

Protein structure determination methods like X-ray crystallography can only provide static snapshots of well-ordered proteins (or parts thereof). The resulting protein structures are often preceded and then augmented by spectroscopic methods like nuclear magnetic resonance (NMR) labeling, EPR or Fourier transform infrared spectroscopy (FTIR), to which intermediate or even transient states are easier accessible than to crystallography. Moreover, these methods can obtain more data on the dynamics of unstructured regions, for example provide inter-label distance measurement with double electron-electron resonance (DEER) or contact probes using florescent labeling techniques. But they mostly lack atomic resolution or provide no detailed structural information. To overcome these limitations, experimentally obtained information is often refined or extended using MD simulations [111]. Partial or fragmented structural data can be assembled into a coherent structural model by applying force-fields and energy minimization techniques [112] that in turn draw from structurally relevant data like inter-label distances [29]. Starting MD simulations based on crystal structures of well-defined states, yields dynamic structural data that can help to interpret experimental data that has lower time- and/or spatial resolution.

1.2.2 *Simulating GPCRs*

For GPCRs, solved structures are available for inactive and active receptor states of various GPCR types and classes [34]. However, intermediate and transition states connecting the inactive and active structures are lacking, hence several approaches employing state of the art MD techniques were used to elucidate the activation mechanism of GPCRs. Adiabatic biased MD (ABMD) simulations suggested putative active states of RhR, including a stabilizing role of K231 [113] and explored the effect of ligand on conformational receptor stability [114]. Specialized computer hardware allowed observing the deactivation pathway of β_2AR, thus probing the G protein-coupled receptor (GPCR) activation mechanism [54]. Using accelerated MD simulations, the activation of muscarinic acetylcholine receptor M2 could

be directly observed [115], predicting some features of the later crystallographically determined M2 structure [33]. Markov state model analysis of two milliseconds of β_2AR dynamics yielded multiple activation pathways which are differentially used by agonists and inverse agonists [116]. Simulations employing a structure-based force-field investigated and compared the activation and deactivation process in β_2AR and RhR [117].

The coupling interface between receptor and G protein in itself has, however, not received the focused attention maintained for the receptor activation process.

1.3 STUDY AIMS

The coupling interface between GPCR and G protein or arrestin is comprised of the interactions created during complex formation. Complex formation requires structural changes in both binding partners, the receptor and the signaling protein. On the receptor side this includes the outward movement and structuring of TM6 together with structural changes to the highly conserved E(D)R$^{3.50}$Y and NP$^{7.50}$xxY motifs, that result in the formation of the open intracellular crevice to which G protein or arrestin can bind [1, 2, 3]. The G proteins contain intrinsically unstructured regions that become structured when interacting with receptor binding crevice. Experimental data shows that the coupling interface is highly dynamic on both sides, until complex formation results in the adoption of specific conformations. On the receptor side, TM6 and ICL3 are the involved structural elements [61, 62]. In the G protein, the C-terminus of the α-subunit becomes structured [68, 31, 32].

Here we investigate the mutual, dynamic relationship between the structured and unstructured elements of the receptor binding crevice and the intrinsically unstructured regions of the signaling proteins. This relationship has functional implications for complex formation, coupling specificity and signal transduction. First, we make the case that the ICL3 in rhodopsin is more dynamic than apparent from crystal structures. We complement spectroscopic experimental data with MD simulations to elucidate the role of the G protein in stabilizing ICL3. Next, we investigate the dynamic TM6 behavior in active β_2AR and its role for G protein coupling specificity. Finally, we turn to the G α5-helix and its mechanistic role in signal transduction from receptor to G protein.

To obtain data on the protein dynamics, we perform molecular dynamics simulations to augment the experimentally determined but static picture of well-ordered protein structures. MD simulations have been successfully applied to membrane proteins and especially to GPCRs (Section 1.2.2) to gain insights into their dynamics. Our simulations run in the order of 200 ns. In this time-scale, it is expected to

observe structural changes like hinge-bending and loop motions that arise from the removal of stabilizing interactions [118].

1.3.1 *Role of ICL3 in complex formation*

Structures of rhodopsin obtained by X-ray crystallography, show the ICL3 loop of rhodopsin and its apoprotein opsin resolved. There, ICL3 exhibits an ordered conformation in both inactive and active receptor states and regardless of whether a $G\alpha_t CT$ peptide is bound or not [39, 30, 31].

Employing MD simulations, we find that the ICL3 becomes unstructured in active and inactive Opsin, only adopting a structured conformation upon binding of $G\alpha_t CT$. These findings corroborate a FTIR study and allow structural interpretation of the spectroscopic data.

It will be proposed and discussed that a mutual structuring of the rhodopsin receptor (ICL3) and transducin G protein can provide a fast and precise signal transfer from rhodopsin to transducin. That the coupling mechanism needs these attributes is implied by the perfect switching where rhodopsin activity increases by more than one billion fold from essentially zero activity after activation by light through photon capturing.

1.3.2 *Role of TM6 in coupling specificity*

A number of GPCRs can activate multiple distinct G proteins [13]. The β_2-adrenoceptor is one of them, coupling to both G_s and G_i [4]. Ligands are known, which can bias the receptors coupling to the G protein and/or the subsequent activation of the G protein [27, 25, 26]. However, the (structural) mechanism by which the active β_2AR* can distinguish between G_s or G_i is unknown.

We employ MD simulations of active β_2AR* with and without G_s or G_i derived binding partners to investigate the dynamics of TM6 — one of the key structural elements of the receptor's G protein binding crevice. In MD simulations, we find that $G\alpha_i CT19$ can stabilize the TM6 of β_2AR* in a RhR*-like conformation. With umbrella sampling calculation, we obtain information about the energy barrier that retards formation of the RhR*-like conformation of β_2AR*.

Comparison with MD simulations of RhR* and RhR*•GαCT suggests a structural mechanism by which GPCRs can distinguish between G_s and $G\alpha_{i/t}$. The extent of TM6 outward tilt changes the size of the receptor binding crevice. Depending on the TM6 outward tilt, the binding crevice is not accessible by bulkier G_s C-terminus but the slimmer $G\alpha_{i/t}$ C-terminus can still bind. Conceivably, this mechanism could be used by ligands that selectively stabilize one of

the TM6 conformations and thus bias the signal towards G_s or G_i in β_2AR*.

1.3.3 *Role of α5 in signal transduction*

The intermediary R*•G[GDP] complex is the first complex of receptor and G protein during receptor triggered G protein activation [44, 78, 83]. The C-terminus of the α5-helix in the Gα subunit has been identified as the main interaction site with R* [70, 71, 72, 73, 74]. During nucleotide exchange in R*•G signaling complexes, the α5-helix of the Gα subunit is displaced by rotation and translation [80, 78, 32]. Here we investigate the role of α5 during transition from the intermediary R*•G[GDP] to the nucleotide-free R*•G[empty] complex.

After modeling the interaction of GαCT with R* in the intermediary R*•G[GDP] complex, we perform MD simulations to test if the α5-helix alone can rotate at the receptor G protein coupling interface. We apply this protocol, using the C-terminus of α5 in lieu of the complete G protein, to two GPCRs and their endogenous G protein: rhodopsin with its G protein transducin and β_2-adrenoceptor with G_s.

The MD simulations of GαCT show switching of the α5-helix as a result of interactions at the R* interface alone. We propose that the α5-helix rotates at the coupling interface of the holocomplex during the transition from the intermediary R*•G[GDP] to the nucleotide-free R*•G[empty] complex. The intermediate R*•G[GDP] complex provides a structural scaffold in which the α5-helix can rotate while the receptor and the G protein stay fixed so that the rotation maximally effect the GDP binding pocket, triggering GDP release.

2 METHODS

2.1 MOLECULAR DYNAMICS SIMULATIONS

At a basic level, the dynamics of proteins arise from atoms that change their position over time through interactions with other atoms. Thus, molecular dynamics (MD) simulations obtain the movement of N interacting atoms at position r over time t by solving Newton's equations of motion simultaneously for each atom r_i (Equation 1).

$$m_i \frac{\partial^2 \mathbf{r}_i}{\partial t^2} = F_i$$

Equation 1: Newton's equation of motion over time t for an atom at position $\mathbf{r}_i$ with mass m_i and force F_i.

The force acting on each atom $\mathbf{r}_i$ is derived from a potential energy function U that includes all interactions in which the atom at position $\mathbf{r}_i$ takes part in, as well as constraints defined for the atom (Equation 2).

$$F_i = -\frac{\partial U}{\partial \mathbf{r}_i}$$

Equation 2: Force F_i acting on atom $\mathbf{r}_i$ derived from the potential energy U.

The movement of the atoms can be followed over time using this framework. Additional effort is needed to ensure that temperature and pressure of the system stay constant or in relevant ranges as required (e. g. by coupling the system to temperature and pressure *baths*). Since the system is finite, the boundaries of it require special treatment (e. g. by employing periodic boundary conditions) to ensure meaningful results. However, when everything is taken care of, the end result of a MD simulation is a trajectory of atomic positions with very high spatial and temporal resolution.

Biological systems

Biologically relevant systems consisting of, for example, proteins and solvent molecules, easily contain in the order of 10^4 to 10^6 atoms. Many biological processes, like signal transduction or protein-protein

binding, occur on timescales ranging from nanoseconds to milliseconds. In order to observe these processes, the system need to be simulated for that long, or longer to get statistically robust data [118]. Generally, simulation times are sought that can be compared to experimental data [52, 104]. Another area where MD simulations have proven useful, often on shorter timescales, is structural modeling and evaluation of general dynamic capabilities of a system [119, 120].

In general, there is a strong tendency that the longer the timescale or the larger the biological systems, the more interesting the simulation results potentially are. The basic equations that drive MD simulations are — not without coincidence — neither complicated nor computationally expensive as such. However, the challenge lies in scaling the computations to work on actual hardware for long timescales and large systems. For that, efficient distributable algorithms and fast implementations for specific processors are required, which immensely complicates performing actual MD simulations. Most simulations run on commodity hardware, but to be able to tackle more computationally demanding problems, there have been efforts to create specialized hardware solely for running MD simulations [121, 122]. The most recent of such projects is the *Anton* machine that also optimizes the distributed hardware architecture for the algorithms needed to perform the simulations [123, 124, 106, 54, 125].

In the following section we will give an overview on how program packages like *GROMACS* work to perform MD simulations. There are many parameters to be set before a simulation can be started and a basic understanding of the inner workings of a simulation package helps to avoid errors when using it to perform simulations and interpreting the resulting trajectories.

2.1.1 *Simulation algorithm*

Every simulation has to start from somewhere (i. e. an initial setup), so an essential part of setting up a simulation is obtaining relevant starting positions for the biological question at hand. Usually, these are positions of the various molecules the simulation system is comprised of, which are biological macromolecules like proteins or DNA/RNA, solvent molecules (often water), ions, lipids and ligand molecules. For these particles, the topology of each molecule is needed in the form of geometric constraints that define what molecule they are. These constraints consist of bonds length, bond angles and dihedral angles. Additionally, the elemental composition of the molecules is required as well as charge, vdW radius and neutral interaction parameters for each atom. Such parameters are usually supplied by so called force fields that in turn derive them from (and fit them to) experimental data. Together, the starting position and the system topology provide the initial conditions, i. e. the input

for the simulation algorithm (Scheme 2). If necessary, initial velocities can also be input, otherwise they are generated, for example, from a Boltzmann distribution at a specific temperature.

- Initial conditions input
- Simulation loop
  - Force calculation
  - Position & velocity update
  - Constraints inclusion
  - Data (positions) output
- Resulting trajectory analysis

Scheme 2: Generic MD simulation algorithm

Given the initial conditions, the system is simulated following Newton's equations of motion in the simulation loop while at the same time considering constraints derived from the system topology (Scheme 2). The simulation loop generally consists of three steps: force calculation, position & velocity update and constraint inclusion. Importantly, the atom positions and other data are written out during the simulation loop, too.

The first step, the force calculation, derives the force $\mathbf{f}_i$ acting on each atom $\mathbf{r}_i$ from a potential energy function U (Equation 2). Non-bonded interactions consist of interactions between neutral and charged pairs of atoms. These interactions are modeled with the Lenard-Jones (Equation 4) and the Coulomb (Equation 5) potential respectively. Intra-molecular, bonded interactions are represented by bond-stretching, angle-bending and dihedral-torsion potentials between pairs, triples and quadruples of atom positions $\mathbf{r}$, respectively (Equation 6). For the second step, based on the forces acting on each atom, the updated positions and velocities can be directly calculated, for example, by using the velocity Verlet algorithm (See Equation 3 and Listing 1). During this step, temperature coupling (to a virtual temperature bath) can be achieved by globally rescaling the velocities of each atom so that the total kinetic energy corresponds to a target temperature.

The inclusion of distance constraints is the third step and used to ensure that the simulated molecules retain their initial topology. MD simulations are only an approximation and for performance reasons it is desirable to increase the simulation time-step as much as possible (to cover more simulated time per unit of computing time). However, bonds, for example, have very high vibrational frequencies

$$\mathbf{p}_i(t + \frac{1}{2}\partial t) = \mathbf{p}_i(t) + \frac{1}{2}\partial t \mathbf{f}_i(t)$$
$$\mathbf{r}_i(t + \partial t) = \mathbf{r}_i(t) + \partial t \mathbf{p}_i(t + \frac{1}{2}\partial t)/m_i$$
$$\mathbf{p}_i(t + \partial t) = \mathbf{p}_i(t + \frac{1}{2}\partial t) + \frac{1}{2}\partial t \mathbf{f}_i(t + \partial t)$$

Equation 3: Velocity Verlet algorithm at time t, with the momentum $\mathbf{p} = m\mathbf{v}$ as the product of mass m and velocity $\mathbf{v}$, the atoms position $\mathbf{r}_i$, force $\mathbf{f}_i$ and time step ∂t. See Listing 1 for a pseudo code implementation of the algorithm.

```
loop{
    p = p + 0.5 * dt * f
    r = r + dt * p / m
    f = force(r)
    p = p + 0.5 * dt * f
}
```

Listing 1: Velocity Verlet pseudo code. The momentum `p = mv` as the product of mass `m` and velocity `v`, the atoms position `r`, force `f`, time step `dt` and a function `force(r)` to caclulate the force acting on `r`.

and would require very small time-steps to faithfully represent their dynamics. Therefore, algorithms like *LINCS* [126] and *SETTLE* [127] have been developed which can guarantee that constraints on bond length in particular and atom to atom distances in general are fulfilled at the end of a simulation time-step. Also performed at this step is pressure coupling (to a virtual *pressure bath*), often by scaling the simulation box and the atom positions so that the system conforms to a target pressure. Pressure and temperature coupling are needed to simulate an NPT ensemble where the number of atoms (N), the pressure (P) and temperature (T) are constant. Finally, the atoms position are written out to create the trajectories that are later used for analysis.

2.1.2 *Force contributions and constraints*

MD simulations must consider a number of non-bonded and bonded (intra-molecular) interactions to sufficiently approximate the dynamic behavior of the simulated atoms over time. However, apart from describing the forces that act on an atom, being able to compute them fast and efficiently is a prime objective when choosing the equation and models to calculate them.

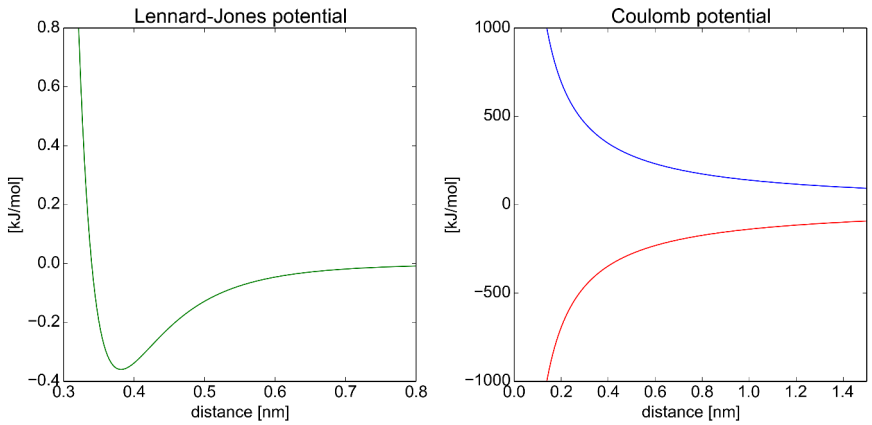

Figure 6: The Lennard-Jones potential (left) is plotted for a carbon-carbon interaction with parameters from the `amber99sb-ildn` force-field [129]. The Coulomb interaction (right) shows the potential for particles with equal elementary charges (blue) and with the inverse sign (red).

Non-bonded interactions

The Lennard-Jones potential describes the neutral (uncharged) interaction between a pair of atoms as a function of their distance. It consists of a weakly attractive term dominant at larger distances and a strongly repulsive term that becomes dominant when the atoms approach each other (Equation 4, Figure 6). The attractive long-range term represents the van der Waals force and the short-range repulsive term approximates the Pauli-repulsion created by overlapping electron orbitals. The Lennard-Jones potential has two parameters, ε and σ, which are specific for the type of interacting atoms. These parameters are created by fits to experimental data and provided by force-fields. Other potentials for describing neutral atom interactions, the Buckingham potential or the Stockmayer potential, provide a more realistic description but are more expensive to compute than the Lennard-Jones potential [128].

$$v^{LJ}(\mathbf{r}_i, \mathbf{r}_j) = 4\varepsilon \left[\left(\frac{\sigma}{d_{ij}} \right)^{12} - \left(\frac{\sigma}{d_{ij}} \right)^{6} \right]$$

Equation 4: The Lennard-Jones potential is used to model the neutral interaction between a pair of atoms. The parameters ε and σ are specific for the types of the interacting atoms r_i and r_j.

The Coulomb potential describes the interaction between a pair of atoms resulting from electrical charges as a function of the distance between the atoms. All atoms in charged or uncharged molecules have individual partial charges resulting in an attractive potential for charges with opposing signs and a repulsive potential for atoms with

the same signs (Equation 5, Figure 6). The parameters Q_i and Q_j are the partial charges of the interacting atoms and are usually calculated for force-fields by solving quantum mechanical equations for smaller parts of whole molecules.

$$v^{\text{Coulomb}}(\mathbf{r}_i, \mathbf{r}_j) = \frac{Q_i Q_j}{4\pi\epsilon_0 d_{ij}}$$

Equation 5: The Coulomb potential is a pair potential between point charges in vacuum with a permittivity of ϵ_0. The parameters Q_i and Q_j are the charges of the interacting atoms r_i and r_j; d_{ij} is their distance from each other.

The non-bonded interactions must be evaluated for all pairs of atoms when applied strictly (Equation 6), which is computationally very expensive as there are $\frac{1}{2}N(N-1)$ such interactions in a system with N atoms. However, a closer look at the interaction potentials suggests that not all atom pairs need to be considered all the time. At increasing distance, the forces exerted by the non-bonded interactions quickly get smaller. The Lennard-Jones potential approaches zero at an exponential rate in the order of d_{ij}^{12} and the Coulomb potential falls off quadratically with respect to the inter-atom distance. To exploit this rapid decline in interaction strength the neighbor-list data structure has been developed. Such a data structure is created for each atom (central atom) and includes all atoms that may interact with the central atom for a given time span. During that time, only the atoms in the neighbor-list are used when calculating the non-bonded interactions for the central atom. After the given time span, the neighbor-list is updated.

$$U_{\text{non-bonded}} = \sum_{i,j}^{\substack{\text{atom}\\ \text{pairs}}} v^{\text{LJ}}(\mathbf{r}_i, \mathbf{r}_j) + v^{\text{Coulomb}}(\mathbf{r}_i, \mathbf{r}_j)$$

Equation 6: The non-bonded interaction potential is the sum of the Coulomb v^{Coulomb} and Lenard-Jones v^{LJ} potentials between all pairs of atoms i, j.

Intra-molecular, bonded interactions

In addition to non-bonded interactions, there are a number of bonded interactions, that characterize the topology of the simulated molecules. These intra-molecular interactions are described by bond-stretching, angle-bending and dihedral-torsion potentials between pairs, triples and quadruples of atom positions, respectively (Equation 7). The actual constraining is performed during the simulation

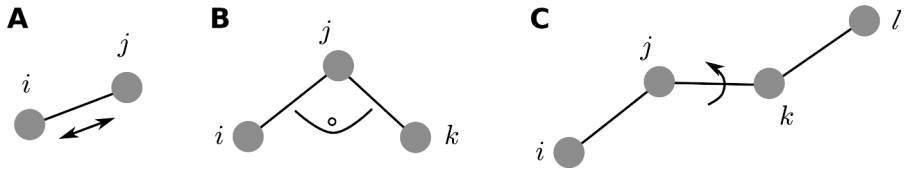

Figure 7: Schematic display of bonded interactions for (A) bond-stretching between pairs of atoms, (B) angle-bending in atom triples and (C) dihedral-torsion in atom quadruples.

loop by algorithms like *LINCS* [126] for general molecules and by *SETTLE* [127] optimized for water molecules. They change the atom positions to guarantee that constrained atom to atom distances are fulfilled at the end of a simulation time-step. Using these distance constraints, the bond length can be constraint but also bond and dihedral angles by expressing the angular constraints with multiple distance constraints.

$$
\begin{aligned}
U_{\text{intra-molecular}} = & \sum_{i,j}^{\text{bonds}} v^{\text{bond-stretching}}(\mathbf{r}_i, \mathbf{r}_j) \\
& + \sum_{i,j,k}^{\substack{\text{bond} \\ \text{angles}}} v^{\text{angle-bending}}(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) \\
& + \sum_{i,j,k,l}^{\substack{\text{dihedral} \\ \text{angles}}} v^{\text{dihedral-torsion}}(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k, \mathbf{r}_l)
\end{aligned}
$$

Equation 7: The intra-molecular, bonded interactions are calculated between pairs, triples and quadruples of atom positions $\mathbf{r}$ representing bond-stretching, angle-bending and dihedral-torsion potentials respectively (Figure 7).

2.1.3 *State of the art*

With enough raw computational power, long timescales become accessible to MD simulations and more complete samplings can be obtained without the need to introduce a bias or sacrifice precision. Progress is fueled by enhancements of hardware and software. The current generation of MD software packages can utilize the massively parallel architecture in graphics processing units (GPUs) [130]. Some effort was made to build special purpose hardware for MD simulations [121, 122], of which the *Anton* machine is probably the most successful achievement [123, 124, 106, 54, 125]. Also, the ubiquitous cloud-based computing has been recently leveraged for MD simulations by performing a large number of short, independent simula-

tions in parallel and combining their results with Markov state models [116].

Also progressing is the development of algorithms that exhibit better scaling behavior on massively parallel computers for the calculation of long range electrostatic interactions than the particle-mesh Ewald (PME) method. Multilevel summation method (MSM) and fast multipole method (FMM) have been shown to be applicable to MD simulations with promising results [131, 132].

2.2 STRUCTURE PREPARATIONS

Before protein structures, for example, from the protein data bank (PDB), can be used in a MD simulations, some refinement steps need to be performed to create a coherent structure. The (raw) structural data from the PDB may be present in a crystallographic or biological assembly that is not correct for the context you want to perform the simulations in. For example, membrane protein structures resolved by X-ray crystallography are not embedded in a lipid bilayer but surrounded by a layer of detergent molecules. Further, the (raw) structure file may contain — apart from detergent molecules — a shell of ions and water molecules that are not particularly relevant for the simulations as they are artificial, owing to the experimental conditions under which the structure was determined, or easily reproduced during the creation of the whole simulation system.

2.2.1 *Protonation states and internal water molecules*

When performing classical MD simulations, correct assignment of the residues protonation state(s) is especially important. For reasons of computational efficiency, protonation state changes are not allowed during the simulation, as the topology of the system cannot change (see Section 2.1). The protonation status of individual residues generally cannot be inferred from the (raw) structural data as its resolution is not high enough to show single hydrogen atoms. But for most of the residues one can assign the protonation status according to their pKa and the pH at which the simulation will be performed. Here, the GROMACS tool `pdb2gmx` was used to automatically assign a default protonation at physiological pH to most residues. However, for some residues there exists experimental evidence that their immediate environment has a different microscopic pH that induces another protonation state. In the following section, such cases are discussed and properly treated. Additionally, the C-termini of the receptor, peptide and G protein molecules are deprotonated (COO^-), whereas the N-termini are fully protonated (NH_3^+).

Internal water molecules are often not resolved due to lack of resolution. Especially membrane proteins are mostly only resolved at modest resolution so that structural details, such as the placement of internal water molecules, are often not adequately inferable. To fill internal empty polar cavities with water molecules we employed the program DOWSER [133]. By calculating the interaction energies of internal water molecules with their surrounding atoms, DOWSER assess if a water molecule makes a significant energetic contribution to the structural stability. In addition to water molecule positions resolved from (raw) structural data, DOWSER tests all positions where a spherical probe, the size of a water molecule, makes contact with in-

ternal protein surfaces. Here we used DOWSER to obtain the internal hydration sites to prepare structure for simulations.

2.2.2 *Receptors and peptides*

Experimental structure determination often requires modifications made to the specimen to make obtaining data possible or to enhance the quality (or resolution) of the results. Such modifications may range from single point mutations over the truncation of N- and C-termini to the replacement of whole specimen regions. For example, to solve the structure of GPCRs by X-ray crystallography, the flexible third intracellular loop (ICL3) is often replaced with a T4-lysozyme to obtain a more rigid structure with potentially more stabilizing contact in the crystal lattice [59]. Still, some regions within the specimen may not be resolved due to high flexibility which results in missing atoms or residues in the structural data. These regions need to be adapted or modeled to get structures suitable for MD simulations.

RhR receptor

MD simulations of inactive rhodopsin (RhR) were based on the structure of chain A in PDB entry 1U19. The two *palmitoyl* chains attached to the residues C322 and C323 were used as well as the retinal ligand covalently attached to K296. All other non-protein atoms were removed from the structure. In inactive rhodopsin, D8 [134] and E122 [134] are likely protonated.

RhR receptor*

For all MD simulations of active rhodopsin (RhR*) based on PDB entry 3PQR, two *palmitoyl* chains were attached to the residues C322 and C323. The covalently attached retinal (to K296) was kept, whereas all other non-protein atoms were removed from the structure. In RhR*, diverging protonation states were suggested for D83 [135, 134] and two residues in the proximity of the Schiff-base, E113 [136] and E122 [134] as well as E134 [137]. All of these are thus set protonated.

The structures of active RhR* lack a large part of the C-terminus (residues 327 to 348, UniProt entry *P02699*), which is not resolved in any other structure but inactive RhR (PDB entry 1U19). However, because removal of this region has been reported not to affect activation of rhodopsin's natural binding partner, the transducin G protein G_t [138], we did not model this part of active RhR*.

Ops receptor*

Opsin is characterized by a ligand-free binding pocket and the structure of active opsin (Ops*, PDB entries 3CAP and 3DQB) shows empty space within the retinal binding pocket. However, it is highly

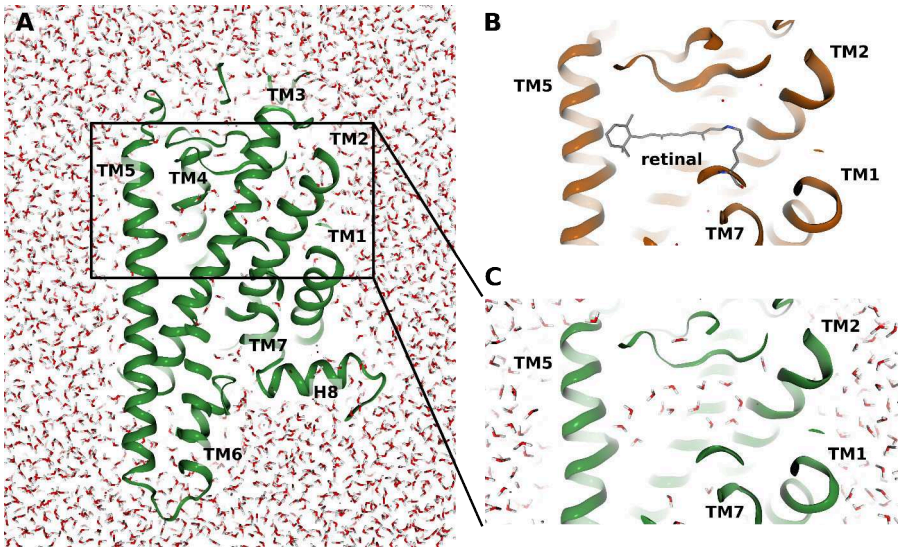

Figure 8: Ops* ligand cavity filled with water. Panel (A) shows Ops* (green) completely embedded in water with no membrane to allow free influx of water through two openings that also allow retinal integration and removal through the lipid bilayer under native conditions [45]. For comparison, the RhR* (orange) ligand binding pocket with covalently attached retinal is shown (B) along a detailed view of the water filled Ops* ligand cavity (C).

unlikely that there is a void inside Ops* as this would lead to a collapse of the receptor. Hence, the retinal binding pocket of Ops* was filled with bulk waters. For that purpose, Ops* was completely embedded in water molecules, and the amino acid side chains and water molecules were allowed to equilibrate during a 60 ns simulation, while the protein backbone atoms were restrained to their initial positions. With this setup, water entered through the two openings that also allow retinal entry and release through the lipid bilayer under native conditions [45]. The steady state was reached after about 20 ns, with approximately 27 water molecules in the (ligand-free) binding pocket and the rest of the protein interior (Figure 8).

As described above for RhR*, the residues D83 [135, 134], E113 [136], E122 [134] and E134 [137] were protonated. Also, the receptor C-terminus is treated as described above for RhR*, that is, not modeled.

RhR•Gα_tCT11/19 and Ops*•Gα_tCT11/19 complexes*

The receptors (RhR*, Ops*) in the complexes (RhR*•Gα_tCT11/19, Ops*•Gα_tCT11/19) were prepared the same as the receptors alone.

For RhR*•Gα_tCT11, the coordinates from the double high-affinity L341, V347 11-mer Gα_tCT peptide variant in complex with RhR* (PDB entry 3PQR) were used. However, the peptide mutations were changed back to their wild-type forms K341 and C347. These mutations were used to aid crystallogenesis by enhancing the interaction

between peptide and receptor. But for the simulations we do not need that stabilizing effect and are more interested in the native contacts. Similarly, for Ops*•Gα_tCT11, the coordinates from the high-affinity L341 peptide variant in complex with RhR* (PDB entry 3PQR) were used. However, the point mutation in the peptide was changed back to its wild-type form K341, to obtain the native form. Therefore, the same 11-mer Gα_tCT (G340–F350) peptide was used for the RhR*• Gα_tCT 11 and Ops*•Gα_tCT 11 complexes.

For simulations of both receptors — RhR* and Ops*, in complex with 19-mer Gα_tCT (332–350) — the 11-mer Gα_tCT was extended N-terminally by 8 amino acids using the geometries of an ideal α-helix.

β_2AR receptor*

The starting conformations for simulating of active β_2-adrenoceptor (β_2AR*) were prepared based on the agonist-bound X-ray structures from co-crystals of β_2AR*•G$\alpha_s\beta\gamma$ (PDB entry 3SN6) [32]. First, the G$\alpha_s\beta\gamma$ was removed from the complex along with the Nb35 nanobody. Non-protein atoms were also removed with the exception of the receptor-bound agonist. The far C-termini of the β_2AR* structure (342–413, UniProt entry *P07550*), not resolved in these complexes, were not modeled as removal of this region has been reported not to affect adenylate cyclase activation [139].

Also — for the simulations of β_2AR*— the T4-lysozyme was removed from the N-terminus and a *palmitoyl* chain was ligated to C341 of β_2AR*. Unresolved atoms from the side chains of residues 63, 97–99, 101, 149, 175, 192–195, 267, 269–272, 299, 301–302, 304, 306 and 333 were added applying the standard geometries from the Dunbrack 2002 library [140]. Three stabilizing mutants (M96T, M98T and N187E) in β_2AR* were changed back to their respective wild-type form, to obtain the native sequence. The coordinates for the missing residues of extra-cellular loop (ECL) 2 (176–178) were taken from the β_2AR*-structure (PDB entry 3PoG) where ECL2 is resolved [60]. The conformation of residues 240 to 264 from ICL3, which are not critical to receptor function [141], were modeled with help of the fragment-based loop modeling program SuperLooper [142].

In β_2AR*, E122 was set protonated because it is in close contact with the hydrophobic lipid tails in the middle of the lipid bilayer (as suggested by Dror et al. [54]).

β_2AR•G$\alpha_s\beta\gamma$ and β_2AR*•Gα_sCT11/19 complexes*

The structure of the active β_2AR* receptor in the β_2AR*•G$\alpha_s\beta\gamma$ and β_2AR*•Gα_sCT11/19 complexes was prepared the same as for the simulations of the receptor alone.

G$\alpha_s\beta\gamma$ was prepared as follows. The missing Gα_s N-terminal residues 1–8 were modeled using standard geometries before a *palmi-*

toyl chain was ligated to residues C3 and G2, each [143]. A *geranylgeranyl* chain was ligated to residue 68 of the $G_s\gamma$-subunit [144], after the missing residues 1–4 of the N-terminus and residues 63–68 of the C-terminus were added. The mutated residues G72S in $G\alpha_s$ and M1Q in the $G_s\beta$ were changed back to the wild-type form. Unresolved atoms from the side chains of residues 24, 35, 58, 59, 94, 118, 136, 139, 188, 189, 191, 194, 195, 201, 216, 240, 300, 322, 369 in $G\alpha_s$, 1, 42, 129, 130, 172 in $G_s\beta$ and 62 in $G_s\gamma$ were added using standard geometries from the Dunbrack 2002 library [140]. The conformation of the missing residues 60–70, 85–87, 203–204 and 256–262 in $G\alpha_s$ were again modeled with SuperLooper [142].

For simulations of 11-mer $G\alpha_s$CT (residues 384–394) and 19-mer $G\alpha_s$CT (residues 376–394), the coordinates from the β_2AR*•$G\alpha_s\beta\gamma$ complex were used by truncating the $G\alpha_s$ subunit N-terminally.

β_2AR•Gα_iCT complex*

An initial β_2AR*•$G\alpha_i$CT19 complex was created from MD simulations based on the crystal structure complexes of β_2AR*•$G\alpha_s\beta\gamma$ (PDB entry 3SN6) and RhR*•$G\alpha_t$CT (PDB entry 3PQR). First, a $G\alpha_i$CT19 starting position within the β_2AR* binding crevice was obtained that resembles $G\alpha_t$CT11 bound to RhR*. To this end, β_2AR* was superposed with the RhR*•$G\alpha_t$CT complex without employing the flexible TM6. Thus, the superposition was guided by a sequence alignment of β_2AR* and RhR* and employed the Cα atoms from TM1–5 and TM7. The position of $G\alpha_t$CT in this superposition was then taken as the starting position of $G\alpha_i$CT, which was created from $G\alpha_t$CT by changing I338 to V. In this starting position, no contacts are formed between TM6 and $G\alpha_i$CT. The peptide is initially attached to ICL2 and TM5 but can move freely within the β_2AR* binding crevice.

RhR•Gα_sCT complex*

The $G\alpha_s$CT starting position for the RhR*•$G\alpha_s$CT complex was obtained following the same sequence alignment and superposition protocol as for β_2AR*•$G\alpha_i$CT but with RhR* as the receptor and $G\alpha_s$CT as the peptide target. Thus, resulting in a complex composed of the RhR* structure and a $G\alpha_s$CT peptide bound in a way resembling $G\alpha_s$CT bound to β_2AR* in the β_2AR*•$G\alpha_s\beta\gamma$ complex.

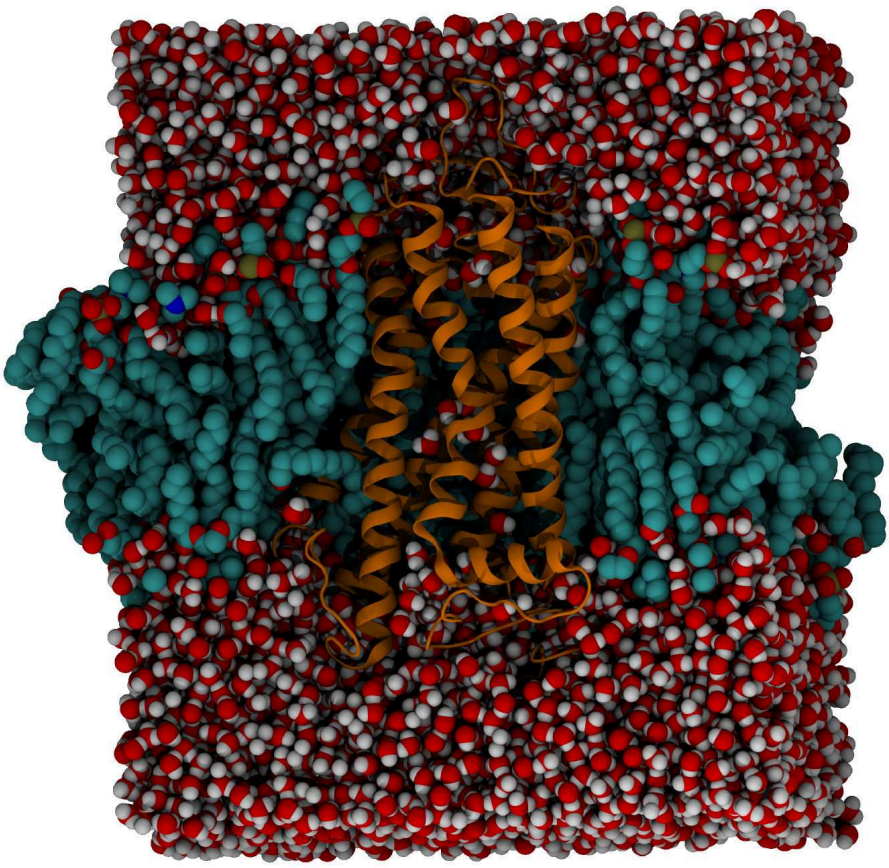

Figure 9: Simulation system with proteins, lipids and water. A rhodopsin molecule (orange cartoon) is embedded in a lipid bilayer surrounded by water molecules (space-filling atoms colored by element, carbon: turquoise, oxygen: red, nitrogen: blue, phosphorus: olive, hydrogen: white). For clarity, lipids and water molecules that would be occluding the view of the receptor are not shown.

2.3 PERFORMING MOLECULAR DYNAMICS SIMULATIONS

To perform the intended classical molecular dynamics simulations of our membrane proteins, the prepared structures (see Section 2.2) need to be put into a suitable water/lipid environment to create a simulation system to be used with the GROMACS simulation package. Further, topology definitions of the prepared structures need to be created with the GROMACS tool `pdb2gmx`. For that, force-field parameters are required that adequately describe the topology and the potential interactions (see Section 2.1). With all the necessary simulation inputs in-place, a preparatory multi-step protocol tries to minimizes the occurrence of artifacts due to the (initially) composed nature of the created simulation system (Figure 9).

2.3.1 *System preparation*

System preparation and subsequent minimization and equilibration were performed with the GROMACS suite (version 4.5) [145]. The prepared (see Section 2.2) receptor and GαCT structures and their complexes were inserted into an equilibrated bilayer of *dimyristoylphosphatidylcholine* (DMPC) using the GROMACS `g_membed` tool [146].[1]

Parameters for the DMPC lipids were derived from Berger et al. [147]. For water, the SPC/E model [148] was used. A salt concentration of 0.15 mol/L was obtained by adding Na^+ and Cl^- ions to the system with the GROMACS tool `genion`. The AMBER99SB-ILDN force field [129] was used for proteins and ions. Ligand parameters for the agonist *5-hydroxy-4H-benzo[1,4]oxazin-3-one* (a.k.a. BI-167107) of β_2AR* were created with the PRODRG2 web-server [149]. Parameters for the *11-cis* retinal in RhR where adapted from Kandt, Schlitter, and Gerwert [150], and parameters for the deprotonated *all-trans* retinal in RhR* were adapted from Mertz, Lu, Brown, and Feller [151].[1]

2.3.2 *Simulation protocol*

To obtain clash-free structures suitable for MD simulations, an energy minimization was performed in GROMACS using the steepest descent algorithm until the maximum force went below 1000 kJ/mol/nm. In the following equilibration step, the energy minimized structure was simulated for 20 ns with all protein backbone atoms restrained to their initial positions. This allows for relaxation at the protein-membrane, protein-water and the membrane-water interfaces so that voids are filled and side chain packing is optimized. For the production MD simulations, the position restraints were lifted.[1]

Based on the equilibrated systems, the production runs were started with different initial velocities obtained from Boltzmann distributions at 320 K. For equilibration and production runs, all bonds were constrained using the LINCS algorithm [126], with the exception of water bonds, which were constrained by the SETTLE algorithm [127]. The temperature was kept constant by coupling the system to a temperature bath of 320 K, which is high enough to keep the DMPC membrane from entering the gel phase. The temperature coupling was performed using the velocity-rescaling thermostat of Bussi et al. [152] with a time constant of 0.2 ps. Long range electrostatics were calculated with the PME method [153]. Berendsen pressure coupling was performed with a time constant of 2.0 ps and semi-isotropic scaling, separating scaling in the membrane plane directions from the

1 paragraph adapted from [56] and/or [101]

z-direction (i.e the membrane plane normal). The integration time step used for all simulations was 0.002 ps.[1]

2.3.3 *Umbrella sampling calculations*

Umbrella sampling (US) facilitates sampling of the conformational space by applying a restraining potential along a transition coordinate. By employing umbrella sampling over a series of windows, ranges on the transition coordinate can be sampled, which would be inaccessible to direct sampling due to energy barriers of the transition coordinate. The resulting series of histograms contains the biased distribution along the transition coordinate. The weighted histogram analysis method (WHAM) is employed to unbias and combine the histograms [154]. From the resulting distribution, the potential of mean force (PMF) can be calculated at temperature T as

$$PMF(c) = -k_B T \cdot \ln p(c)$$

for the probability p of the transition coordinate c and with the Boltzmann constant k_B [154].[1]

Here, the transition coordinate for the free energy calculations of TM6 inward movement and β_2AR* Gα_iCT interaction was selected from the trajectories of a series of β_2AR*•Gα_iCT19 MD simulations (Figure A2, simulation 8). Along the selected trajectory, umbrella sampling MD simulations were performed with respect to the TM6 inward transition by applying the umbrella potential to the upper part of TM6, namely to the backbone atoms of residues 265 to 277. We simulated 36 US windows for 200 ns each. WHAM was then employed to obtain the PMFs from the last 100 ns from each US window and the error was estimated by the standard deviation of a block-wise (three equally sized blocks) analysis.[1]

1 paragraph adapted from [56] and/or [101]

2.4 ANALYSES OF MD SIMULATIONS

The primary result of MD simulations are atom trajectories which provide the position of each atom in space and time. Using molecular visualization tools, the trajectories can be interactively viewed according to the topology of the simulated system, that is, which atoms are connected/bonded to form molecules. Such visual exploration and analysis immediately gives an overview of the general behavior of the simulated molecules. However, to obtain summarizing descriptions of the systems behavior and intrinsic properties that do not become apparent from the atom positions alone, dedicated calculations are performed.

2.4.1 *Secondary structure assignment*

During MD simulations, the secondary structure of the simulated proteins may change. Most of the simulated systems are based on crystal structures where parts of the proteins are stabilized by crystal lattice contacts or by crystallization helper molecules. Therefore an accurate secondary structure assignment cannot be obtained from the crystal structure but must be obtained dynamically from the simulation data. In our secondary structure analyses, we focused on the α-helicity of the ICL3, TM6 and the GαCT. The ICL3 and TM6 show extended crystal lattice contacts (RhR*) or are only partially resolved (β_2AR*). The GαCT peptides are derived by N-terminal truncation from longer α-helical structures and may show increased N-terminal flexibility.

In this study, we use the DSSP [155] tool together with the GROMACS script `do_dssp` to calculate the secondary structure of each residue over time. DSSP differentiates between eight structural types and assigns each residue a single type. The structural types can be classified into three helical types (α-helix, π-helix and 3_{10} helix), two strand types (β-sheet and β-bridge) and three unordered types (turn, coil and bend).

2.4.2 *RMSD and RMSF calculations*

MD simulation data provides a position for every atom at any time. First steps to summarize this wealth of data are measures that subsume the positional data either over all atoms or over time by calculating the root mean square deviation (RMSD) or root mean square fluctuation (RMSF) respectively. GROMACS provides `g_rms` and `g_rmsf` tools to efficiently calculate RMSD and RMSF values.

The RMSD is a measure for the average distance between pairs of atoms (Equation 8). By calculating the RMSD between, for example, the starting structure of a simulation (or any other reference structure) and the structure at any given time frame, the deviation from

the starting (reference) structure can be plotted over time. To focus on the overall structural changes, the RMSD is usually calculated for Cα or backbone atoms, thus ignoring the side-chain fluctuations. Additionally, before comparing the structures by applying the RMSD formula, the structures are superposed so that the RMSD reflects the internal differences and not the global rotation or translation.

$$\mathrm{RMSD}(x, y) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} |x_i - y_i|^2}$$

Equation 8: RMSD equation. N is the number of atom positions in each x and y.

The RMSF is the standard deviation of atomic positions over the simulation time (Equation 9). By calculating the RMSF for each atom, flexible and structured regions become easily distinguishable. As with the RMSD, a preceding superposition is generally necessary. The RMSF provides a structure-wide summary of the atom's dynamic behavior over time. As such, it is well suited to be mapped on a depiction of a (representative) structure to show its dynamic behavior over time in a structural context.

$$\mathrm{RMSF}(x_i) = \sqrt{\frac{1}{T} \sum_{j=1}^{T} |x_i(t_j) - \bar{x}_i|^2}$$

Equation 9: RMSF equation. T is the number of timesteps, $x_i(t_j)$ is the position of atom x_i at time t_j and $\bar{x}_i$ is the time averaged position of x_i.

2.4.3 *Structural superposition*

Most structural measures for analyzing simulation data are intended to provide internal changes and not global rotations or translations. For example, a membrane protein moving in the lipid bilayer may provide interesting diffusional characteristics, but when analyzing the tilt of a TM helix within its helix bundle such motions are all-consuming noise. Therefore, a superposition should be performed beforehand minimizing the RMSD between two structures, or more explicitly between pairs of atoms. The fitting procedure first removes the translational differences by aligning the two structures and then finds the optimal rotation. Two equivalent solutions for finding the optimal rotation have been described [156, 157].

Most of the GROMACS analysis tools provide the option to perform a superposition step before the actual analysis and accept a

group of atoms to fit. Unless otherwise stated, the analyses described here were performed on structures superposed to the initial structure of the respective simulation. Specifically, by superposing the backbone atoms of the transmembrane helices 1–7.

2.4.4 *TM6 position/tilt*

In GPCR activation, the release of TM6 from the 7-TM bundle is a defining moment that leads to the formation of an open intracellular binding crevice. This outward tilt of TM6 can be measured as the distance between the tip of TM6 and the tip of either TM2 or TM3 at the opposing side of the helical bundle. Here we chose TM2, because RMSD calculations of the tips of TM2 and TM3 have shown that TM2 generally moves fluctuates less. So the distance between TM2 and 6 ($d_{TM2–6}$) was used as an indicator of the TM6 tilt. Specifically, it is measured as the distance between the geometric centers for intracellular sections of TM2 and TM6. For TM2, we used the backbone atom positions of the residues 71–75 (RhR*) and 67–71 (β_2AR*); for TM6 residues 244–248 and 265–269, respectively.

2.4.5 *MD cluster analysis*

The structural variability of ICL3 (residues 224–251) in RhR, RhR* and RhR*•Gα_tCT was measured by cluster analyses performed with the GROMACS tool `g_cluster`. The MD trajectories were prepared by a superposition to align all frames on the backbone atoms of the transmembrane helix segments for the inactive and the active receptor with or without GαCT peptide. Cluster analyses were then performed on the aligned trajectories using the `gromos` method [158] and a 2 Å root mean square deviation (RMSD) cut-off to decide when two structures (e. g. ICL3 backbone atoms at two points in time) are neighbors. Shortly, the `gromos` clustering method starts by counting how many neighbors each structure has that satisfy the RMSD threshold. The structure with the most neighbors is taken — including its neighbors — to form the largest cluster. Then, this process is repeated, ignoring all structures that are already in a cluster, until no more structures are left. This results in a series of distinct clusters of structures where each cluster tends to be centered around a representative structure.

2.4.6 *Calculation of peptide helix axis, tilt and rotation*

The first turn of both peptides from the Gα C-terminus (GαCT) is highly dynamic because of the absence of N- terminal intra-helical hydrogen-bonding partners stabilizing this part of the α-helix in α5 (Figure A9). The overall stability of the α-helical conformation

of residues 3–8 and of the turn structure of residues 9–11 was assessed by secondary structure content (Figure A9) using the program DSSP [155]. Based on the α-helical residues, the representative helix axes for the two GαCT peptides can be determined and thus variations in tilt and rotation of the helices (Figure A13). The rotation and tilt movements of $G\alpha_tCT$ and $G\alpha_sCT$ were calculated from the local axis of their α-helical sections as implemented in the GROMACS tool `g_helixorient`. The program calculates the local rotation and tilt for groups of four consecutive Cα atoms. We calculated these values for all α-helical sections and subsumed the local tilt and rotation respectively to get an overall measure of the rotation and tilt movement. The tilt angle between GαCT and the membrane plane is subsumed analogously as the angle between the local helix axis and its projection onto the membrane plane. The membrane plane was defined from the position of the active receptors (R*) relative to the initial lipid bilayer into which R* was inserted and was thus not dynamically recalculated during the MD simulations. Sample calculations of the membrane planes as the best fitting (least square) planes through all lipid head group phosphates confirmed that the orientation of R* within the lipid bilayer does not change much.[1]

2.4.7 *Calculation of GαCT RMSD and lateral motion within the binding cavity of R**

Two different types of backbone RMSD were calculated with the GROMACS tool `g_rms`. The binding cavity RMSD describes the structural integrity of the binding cavity and was calculated as the RMSD of the backbone atoms from the binding cavity after a least-square superposition of the backbone atoms of the binding cavity to the backbone atoms of the binding cavity from the equilibrated system. The binding cavity was defined as all residues of R* that are in contact (atoms within 4 Å) with residues from the C-terminal reverse turn of $G\alpha_sCT$ (residues 391–393) and $G\alpha_tCT$ (348–350). The peptide RMSD describes the movement of GαCT relative to its position in the crystal structure. It was calculated as the RMSD of the backbone atoms from GαCT after the same superposition was used to determine the binding cavity RMSD. Because the N-terminal turn of $G\alpha_sCT$ is very flexible (for DSSP analysis see Figure A9), the first two residues at the N-terminus were omitted for RMSD calculations. The lateral (XY) motion of the C-terminal reverse turn within the membrane plane was calculated from the geometric center of the backbone atoms of the last seven C-terminal residues of GαCT.[1]

1 paragraph adapted from [56] and/or [101]

2.4.8 *Calculation of hydrogen bond and cation-π interaction energies*

Potential hydrogen bonds between $R135^{3.50}$ of RhR*/Ops* and C347.O of GαCT were monitored as a function of the distance between the acceptor atom C347.O and the hydrogen atoms: HE, HH11, HH12, HH21 and HH22 of $R135^{3.50}$ (as named in the AMBER99SB-ILDN force field). The hydrogen bond energies were calculated according to the formula of Espinosa, Molins, and Lecomte [159] as a function of the distance between the acceptor atom and the hydrogen of the donor atom:

$$-0.5 \cdot (50 \cdot 1.1 \cdot 10^3 \cdot e^{-36} \cdot d),$$

where d is the distance between acceptor and hydrogen.[1]

The hydrogen bond interaction energy was calculated as the sum of the hydrogen bonding energies of all acceptor-donor pairs. The hydrogen bonds that include the atoms HE and HH21 of $R135^{3.50}$ contribute most to the total energy. Cation-π interactions between $R131^{3.50}$ of β_2AR* and Y391 of $G\alpha_s$CT were detected with the program CaPTURE [160]. This program also reports an electrostatic interaction energy estimate that agrees qualitatively with *ab initio* calculations at the HF/6–31G** level, but underestimates the magnitude by roughly a factor of two which is about the same as the van der Waals interaction energy that is also reported by this program [160]. Here we use the electrostatic interaction energy as reported by CaPTURE to qualitatively capture the trend of the cation-π interaction in the MD simulations.[1]

2.4.9 *Solvent accessible surface, interaction area and water contacts*

The interface between receptor and GαCT was investigated to monitor the expulsion of water molecules from the interface and to measure the resulting formation of a hydrophobic patch between ICL3/TM6 and GαCT. The solvent accessible surface area (SAS) can be calculated for complete molecular complexes or parts of them with the GROMACS tool `g_sas`. Additionally, for MD simulations that include explicit water molecules, the water accessibility can be obtained directly by counting contacts between specific residues and the entirety of water molecules (i. e. the solvent).

The interaction area between two molecules (i. e. binding partner A and B) can be calculated as the combination of multiple SAS. First, the SAS of the complex (SAS_{AB}) and the SAS of binding partner B (SAS_B) are calculated separately. The difference between the area contribution of B to SAS_{AB} and SAS_B constitutes the interaction area. Further, the fraction between the interaction area and SAS_B can also

1 paragraph adapted from [56] and/or [101]

be calculated. A fraction of one means a binding partner B is completely accessible, whereas a completely buried binding partner has a fraction of zero.

2.5 CLUSTER ANALYSIS OF DOCKING CALCULATIONS

To detect feasible interactions between $G\alpha_s CT$ and the $\beta_2 AR^*$ cytoplasmic crevice, docking calculations were performed by P. Hildebrand following the flexible docking protocol as described in his previous analysis [78] and implemented in the GOLD package [161]. To group similar docking poses from results of 11 independent docking runs of GOLD, an RMSD based cluster analysis were performed. To realize this analysis we applied the single-linkage clustering method with a cut-off of 1.5 Å as implemented in the GROMACS tool `g_cluster`.

2.6 ANALYSIS AUTOMATION STRATEGIES

To facilitate analysis of a large number of different simulated systems, each with many individual simulations (Section 2.2, Table A1), a flexible analysis automation strategy was devised. A number of observations made during initial/previous analyses provided necessary insights to create a flexible protocol. Analyses like "what are the hydrogen bonds between receptor and peptide" need to be performed on many different systems that contain a receptor and a peptide. Generalized, this means applying a tool to various different systems. While doing so, tool-specific parameters like the "maximum hydrogen bond length" do not get changed, but system-specific parameters like the "trajectory input file" do change. Moreover, to comprehensively analyze even a single system, for example MIIb•Gα_tCT (`MIIb_GtaCT11`), many different tools may be used. Following the (just introduced) notion of tool- and system-specific parameters, it is clear that system-specific parameters are fixed while tool-specific parameters change with every tool used.

Now, it would be advantageous to exploit the distinguishing property of tool- and system-specific parameters, that is the observation that they are independent. For that, a *structured project definition* (Section 2.6.1, Listing 2) was created, where tool and system-specific parameters are saved independently from each other. Given such a clear separation, the idea is to flexibly apply tools to systems using their respective parameters. To do so, a *common tool interface* (Section 2.6.2, Listing 3) is needed that understands the tool- and system-specific parameters regardless of what specific calculation an underlying tool performs.

2.6.1 *Structured project definition*

The *structured project definition* is written in the JavaScript object notation (JSON) format and aims to provide a succinct description of all required tool- and system-specific parameters. A simplified excerpt of a *structured project definition* file is given in Listing 2 and the line numbers in this section refer to that listing. Mainly, it is a loosely hierarchical format in which data at lower levels of the hierarchy try to utilize data from the higher levels whenever possible. For example, to get the trajectory file name in line 15, the values of the `__dir__` entries in lines 3, 7 and 14 are concatenated to yield `/home/arose/projects/rho/3pqr/GaCT/all/md_all.xtc`. This helps to minimize repetition of file name parameters.

```
1     {
2         "project": {
3             "__dir__": "/home/arose/projects"  // directory
4         },
5         "systems": {
6             "MIIb_GtaCT11": {
```

```
7                 "__dir__": "rho/3pqr/GaCT/",  // directory
8                 "__variables__": {
9                     // data commonly needed as tool input
10                    // e.g. residue or atom selections
11                },
12                "__sub__":{
13                    "all": {  // concatenated simulations
14                        "__dir__": "all/",
15                        "traj_file": "md_all.xtc",
16                        "part_length": [ 200, 200, 200, 200, 200 ]
17                    },
18                    "md01": {  // individual simulation
19                        "__dir__": "md01/analysis/",
20                        "traj_file": "md01_mc_fit_protein.xtc"
21                    },
22                    // ... 4 more simulations
23                }
24            }
25        },
26        "tools": {
27            "hbond_peptide_receptor": {
28                "__doc__": "Calculate a receptor's internal hbonds",
29                "__name__": "mdkit.hbond.Hbond",
30                "radius": 0.36,  // max hbond distance
31                "output_dir": "hbond/intra_receptor/",
32                "ndx_group1": "Receptor",  // resdiue selection 1
33                "ndx_group2": "Peptide"    // resdiue selection 2
34            }
35        },
36        "defaults": {
37            "gro_file": "md01_protein.gro",  // common input files
38            "ndx_file": "index2.ndx",
39            "figsize": [ 8, 4 ],  // default plot size
40        }
41    }
```

Listing 2: Simplified excerpt of a *structured project definition* (see the running text for a description of its content). The *structured project definition* uses the JSON format[1] to save its content. The JSON format knows a succinct set of values. These are objects (key value pairs seperated by commas enclosed in curly brackets, e. g. `{"key1": value1, "key2": value2}`), arrays (values seperated by commas and enclosed by square brackets, e. g. `[value1, value2]`), strings (enclosed by quotation-marks e. g. `"string1"`), numbers (e. g. `1.23`) or booleans (`true` or `false`). Here we additionally use `//` to denote comments.

The top level of the format consists of the four entries `project`, `systems`, `tools` and `defaults`. The `project` entry contains general data on the project such as its base path (line 2). The `systems` entry lists all different simulated MD systems and their specific parameters, for example the `MIIb_GtaCT11` starting in line 6. Each system entry has a `__sub__` entry containing parameters relevant for any trajectory file that may need to be analyzed. Essential for this is the `traj_file` entry (line 15) and for concatenated trajectory files the length of the individual parts is often useful (line 16). All tools eventually required for analysis are held in the `tools` entry. The example starting at line 27 is the `hbond_peptide_receptor` tool used to "Calculate a receptor's

internal hbonds", as stated in its documentation string (`__doc__`, line 28). The `__name__` entry (line 29) directly refers to the tool class `Hbond` (Listing 3), contained in the `hbond` module within the `mdkit` package comprising all tools. The next entries specify the maximum distance between hydrogen bond partners (line 30), the path where the output files are written to (line 31) and the two residue groups between which hydrogen bonds are to be analyzed (line 32–33).

Flexible analysis framework

The top level organization of the *structured project definition* immediately reflects the separation of the system- and tool-specific parameters that spurred the creation of the format. Moreover, the way of referring to the code class to be executed (`__name__`, line 29), is identical to the module organization of the Python language, which was used to implement the tool framework. Therefore the *structured project definition* can be seen as a *data-driven* (or *parameter-driven*) extension of the Python language, providing domain specific functionality. In other words, the *structured project definition* is as a configuration file for a flexible analysis framework.

2.6.2 *Common tool interface*

This is mostly a broad overview of the analysis framework built upon the *structured project definition* and the *common tool interface*. The framework is implemented in the Python programming language. To start, some detailed information is appropriate to help illustrate how the `Hbond` class example (Listing 3) fits into the context of the framework. The heart of the *common tool interface* is the `Tool` class, which implements the always required boiler-plate code. This makes it possible to easily start new tools like a hydrogen bond analysis tool by simply creating an `Hbond` class that inherits from the `Tool class` (Listing 3). Similar code reuse is achieved by means of the supplemental classes `GromacsMixin` and `PlotMixin`, which provide basic support for working with GROMACS command line tools and for graphical outputs, respectively.

Input & output. Essential when implementing tools following the *common tool interface* definition is setting the `args` & `out` properties and implementing the `calculate` & `analyze` methods. The `args` property defines all parameters that may be set by data from a *structured project definition*. When creating a new class it is initially filled with common `args` from the base classes the new class inherits from, for instance, the `Tool`, `GromacsMixin` or `PlotMixin` classes. The `out` property lists all the files a tool may create. This is useful when the tools are programmatically used and nested; that is, when one tool calls another tool. By clearly defining the input and output of the tools,

```
1     class Hbond( Tool, GromacsMixin, PlotMixin ):
2         args = [
3             # input files and parameters
4             # special to the Hbond tool
5         ]
6         out = [
7             # output files generated by the Hbond tool
8         ]
9         def calculate( self ):
10            # code neccessary to calculate hydrogen bonds
11        def analyze( self ):
12            # creation of plots and lists
```

Listing 3: The stub of the `Hbond` class conforming to the *common tool interface*. The definition in line 1 shows that the `Hbond` class inherits common methods from the `Tool`, the `GromacsMixin` and the `PlotMixin` class. By inheriting from the `Tool` class, setting the `args` & `out` properties (line 2 & 6) and implementing the `calculate` & `analyze` methods (line 9 & 11), the `Hbond` class conforms to the *common tool interface*.

there is no need to know their internal functioning to be able to interoperate with them as modular components.

Computation. The tool's actual functionality needs to be implemented in the `calculate` & `analyze` methods. Both methods may produce the output files expected of the tool. Separation of the methods into two is founded in the type of data processed. To analyze MD simulations, the tools mainly deal with atom coordinate trajectories. The common processing scheme is to first *calculate* some measure from the trajectory data and then *analyze* the extracted values. This process is reflected by the `calculate` and the `analyze` methods. For the Hbond tool, the former method calculates potential hydrogen bonds and records the time and position of the involved atoms. The `analyze` method then, for instance, filters out very infrequent hydrogen bonds, groups the remaining hydrogen bonds by residue and plots their occurrences over time.

Conclusion

The described *common tool interface* makes creation of new tools very easy, with hardly any need for (repetitive) boiler-plate code. This renders it feasible to write simple wrapper tools around the command line tools supplied by GROMACS itself, for example around `g_hbond` to create the Hbond tool. These wrapper tools then re-use highly efficient functionality but adhere to the *common tool interface* so they can be used in conjunction with the *structured project definition*. All the analyses described in Section 2.4 were performed with similarly

created tools that augment existing programs with project-specific post-processing and visualization steps. Hence, the analyses of the MD simulations were automated utilizing the *common tool interface* and beyond that the *structured project definition*.

3

RESULTS

3.1 ROLE OF ICL3 IN COMPLEX FORMATION OF RHR*

GPCR structures obtained by X-ray crystallography show the ICL3 loop of opsin/rhodopsin in an well-ordered conformation in both inactive and active receptor states and regardless of being bound to a $G\alpha_tCT$ peptide or not. To characterize the dynamics of ICL3 in RhR, Ops* and Ops*•$G\alpha_tCT$ (as well as in RhR* and RhR*•$G\alpha_tCT$), we performed MD simulations of the receptors in a lipid/water environment that is more natural than the crystal lattice in which the structures were resolved. Accounting for the experimental conditions is especially important for ICL3, which is involved in crystal lattice contacts that stabilize a well defined conformation that may be artificial. Additionally, we specified the effect $G\alpha_tCT$ on the ICL3 structure by removing $G\alpha_tCT$ from the receptor. The observed ICL3 flexibility is then investigated by RMSD, RMSF and cluster analyses of the simulation data. In the discussion, these MD results will complement data from spectroscopic experiments on the G protein peptide's role in stabilizing ICL3.

3.1.1 *MD Simulation of RhR, Ops* and Ops*•$G\alpha_tCT$ states*

The MD simulations of inactive RhR [39], active Ops* and Ops*• $G\alpha_tCT$ [31] show how the ICL3 structure behaves when taken out of the crystal lattice and embedded into a lipid bilayer within a water/ion solution. By looking at the backbone RMSD values of the ICL3 and the other intracellular loops ICL1 and ICL2, we can establish how much the loops diverge from their initial structure during the simulation. A larger RMSD value indicates either a more flexible structure or that a new stable conformation is adopted and bearing the latter in mind the RMSD can thus be used as a first proxy for flexibility. The analysis shows that in all RhR, Ops* and Ops*•$G\alpha_tCT$ states, the ICL3 is considerably more flexible than ICL1 and ICL2 (Figure 10), suggesting that regardless of receptor activation ICL3 behaves differently from the two other intracellular loops.

RMSD & Cluster analysis

The RMSD analysis (Section 2.4.2) by itself can not detect whether a new conformation is adopted and can not tell if these are frequently occupied and for how long, in other words whether they form distinct conformational sub-states. Frequently occupied confor-

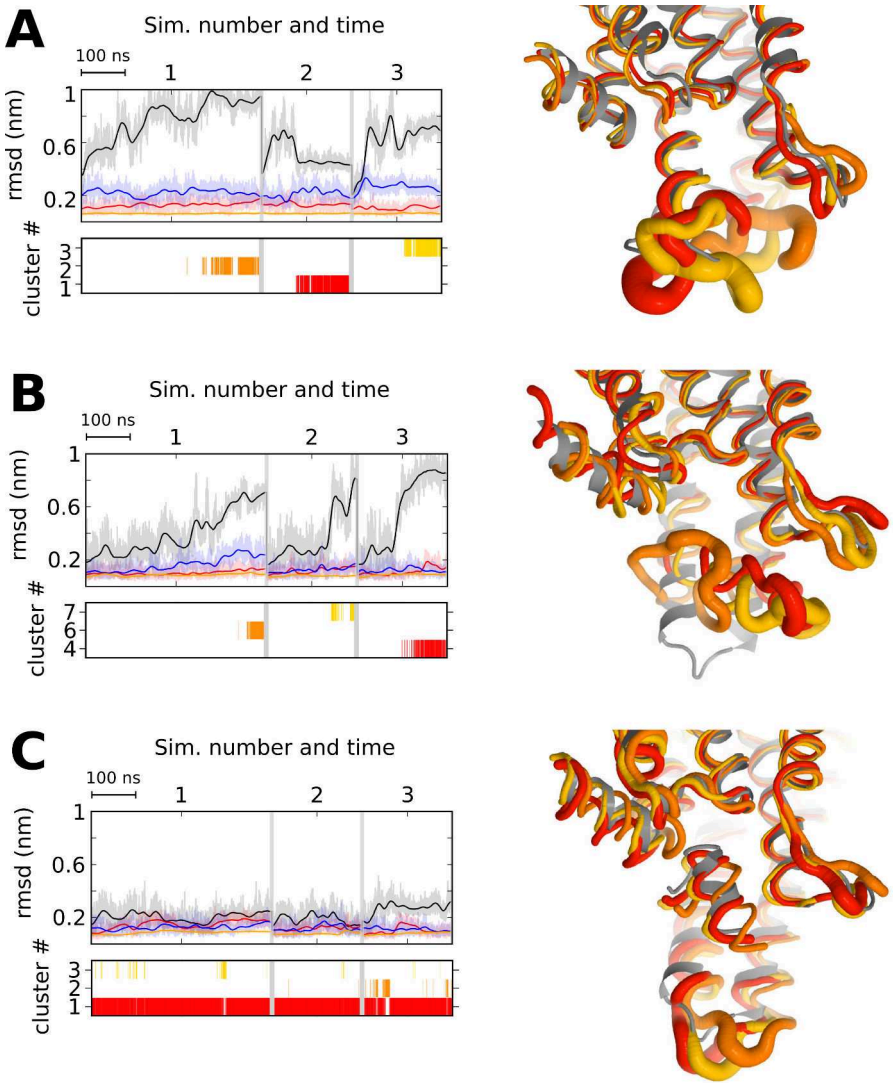

Figure 10: Conformational flexibility of ICL3 in RhR, Ops* and Ops*•Gα_tCT analyzed by MD simulations. (A) Inactive RhR, (B) active Ops*, and (C) active Ops* in complex with Gα_tCT. The bottom row shows for each system the respective X-ray structure (gray) and three representative conformations (red, orange, yellow) obtained by cluster analysis. The tube thickness directly relates to fluctuations of each residue (RMSF) within a given cluster. The occurrences of depicted cluster conformations are given in the middle row. The top row depicts the RMSD for the transmembrane helices (orange), ICL1 (red), ICL2 (blue), and ICL3 (black). (A) The three most occupied conformations (clusters 1–3) are shown, indicating a highly flexible ICL3. (B) Depicted are three metastable conformations (clusters 4, 6 and 7) which are adopted after leaving the initial conformation given by the X-ray structure. It is observed that these conformations vary significantly among each other and with respect to the starting structure. (C) The three most occupied conformations (clusters 1–3) are very similar to the initial conformation taken from the crystal structure. See Figure 11 for the full cluster analysis. *The figure and legend are reproduced from [56]*

mations of ICL3 over the course of all simulations were therefore identified by a cluster analysis of ICL3, namely the residues 224–251 (Figure 11). Here we use representative structures of the three most occupied ICL3 conformations for each of the RhR, Ops* and Ops*•$G\alpha_t$CT based simulations to assess the structural heterogeneity of ICL3 (Figure 10). Comparing representative ICL3 structures from the cluster analysis with the respective crystal structures shows whether the ICL3 structure diverged in the MD simulations (Figure 10).

Cluster analysis of the simulations of inactive RhR shows a high ICL3 flexibility (Figure 10A) which fits well to the high B-factor values found in the crystal structures of inactive RhR (e. g. 1U19, 1GZM). Moreover, while the loop is flexible during the simulations, the TM6 still retains its solvent exposed helical extension as seen in the crystal structures. The RMSD also gives a first indication of a different ICL3 flexibility in the active receptor Ops* compared to Ops*•$G\alpha_t$CT, (i. e. the ICL3 flexibility depends on the presence of the $G\alpha_t$CT peptide. In active Ops*, the ICL3 departs from its initially structured conformation and becomes highly flexible and more and more disordered after > 200 ns (Figure 10B). Also note that the various disordered conformations have lifetimes of about $\sim 10^{-8}$ s (as estimated from the time-lines in Figure 11) and are populated for less than 10% of the respective simulations.

Cluster and RMSD analysis of the simulations of the Ops*•$G\alpha_t$CT complex reveal a rigid ICL3 with its secondary structure preserved (Figure 10C). The largest cluster contains almost 95% of all observed ICL3 conformations (Figure 11C). this cluster shows some minor flexibility, but fluctuates tightly around the mostly α-helical conformation resolved in the crystal structure (Figure 10C, gray). Analyzing the interaction area between receptor and peptide in comparable simulations of the RhR*•$G\alpha_t$CT complex, we observe that the peptide residues L344, C347, G348 and L349 are buried within the receptors intracellular crevice (Figure A8B). In conjunction with L226, V230, V250 and V254 from ICL3, a hydrophobic patch is formed, stabilizing ICL3 in a α-helical conformation. Additionally, hydrogen bonds from K341 in $G\alpha_t$CT to T242 and T243 in the receptor are observed (Figure 17C,F).

RMSF analysis

To obtain a more fine-grained description of the ICL3 flexibility, we calculated the RMSF from the MD trajectory for each amino acid (Section 2.4.2). The RMSD gives a time-resolved description of the whole structure or of individual parts such as the ICL3. Similarly, the cluster analysis is limited to the whole structure or individual parts. In contrast, the RMSF summarizes the flexibility during the complete trajectory in a single value. To compare the ICL3 flexibility of the RhR and Ops* simulations with those of Ops*•$G\alpha_t$CT, we calculated

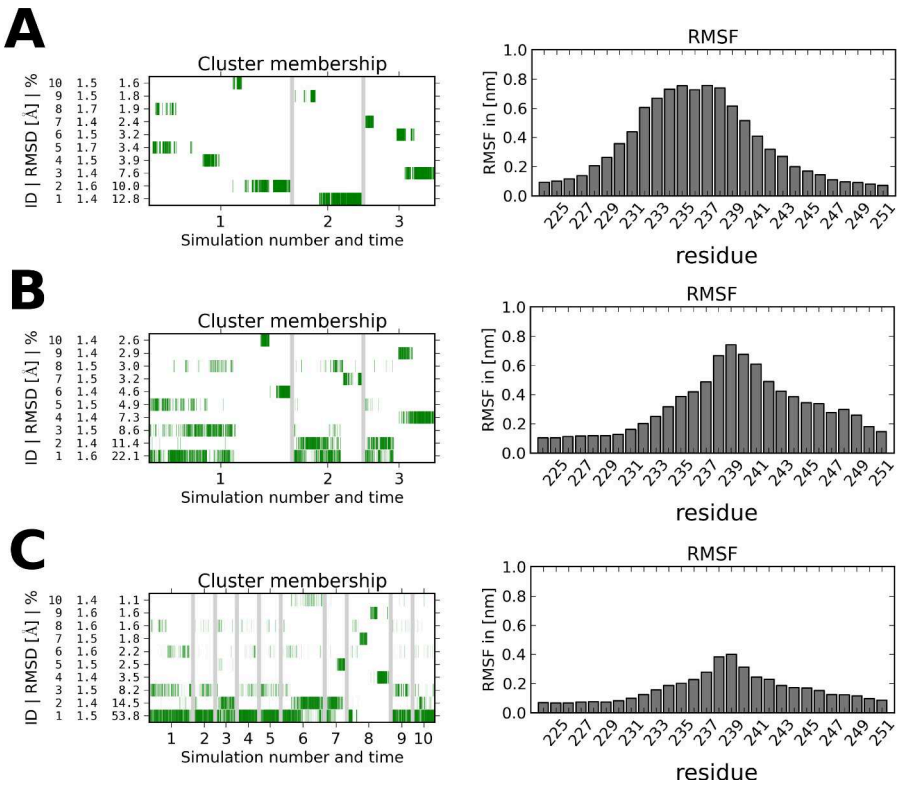

Figure 11: Cluster membership time-line and RMSF plots for (A) inactive RhR, (B) the active conformation Ops* and (C) the active conformation Ops* complexed with $G\alpha_t$CT peptide. For each system, one 400 ns (#1) and two 200 ns simulations (#2 and #3) were run. The cluster membership at each simulation time frame (green) for the 10 biggest clusters of each system is shown in the top row. Each cluster is labeled with its ID, the average backbone RMSD among the cluster members and the percentage of frames within the cluster. The bottom row shows the backbone RMSF of residues 224 to 251. *The figure and legend are mostly reproduced from [56]*

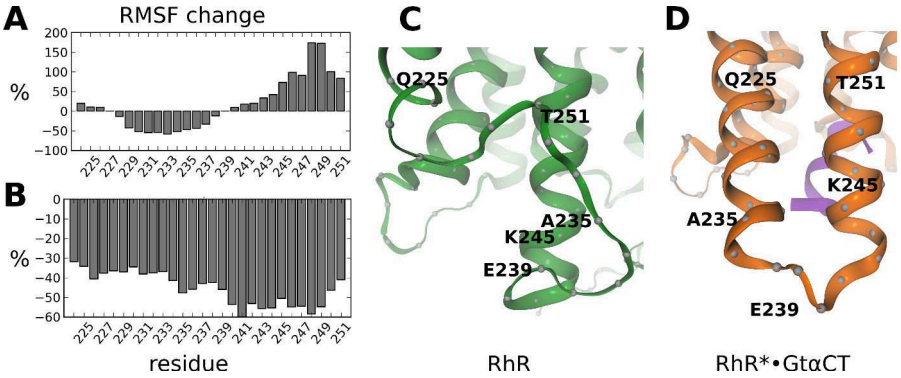

Figure 12: RMSF change between dark state and active conformation (A) and between active conformation with and without $G\alpha_t$CT peptide (B). For comparison, a detailed view of ICL3 in the crystal structures of RhR (C, green, 1U19) and RhR*•$G\alpha_t$CT (D, orange, 3DQB) is shown. Selected residues are labeled at the height of their Cα atom (gray spheres). The $G\alpha_t$CT peptide is colored violet. *The panels (A), (B) and their legends are reproduced from [56]*

how much the RMSF changes (Figure 12). A comparison of RhR and Ops* shows that activation leads to a loss of flexibility particularly in TM5 around A233 but a gain of flexibility in TM6 around K248. The binding of the $G\alpha_t$CT peptide strongly restrains the ICL3 residues as apparent from the comparison of Ops* and Ops*•$G\alpha_t$CT RMSF values (Figure 12). Taken together, these analyses of the simulation data suggest that the ICL3 is an intrinsically unstructured region that becomes structured and more defined when $G\alpha_t$CT binds.

ICL3 dynamics characterization & comparison with experimental data

To summarize, ICL3 is the only of the three intracellular loops (Figure 11, top row) that departs much from the starting conformations taken from the crystal structures as indicated by the large RMSD increase of 5 Å to 10 Å. ICL1 and ICL2 remain near their starting conformations. The addition of the $G\alpha_t$CT peptide, however, dramatically reduces the flexibility of ICL3, which remains structured in all simulations of Ops*•$G\alpha_t$CT.

The ICL3 conformations observed in the simulations of RhR and Ops* built up an ensemble of fluctuations between similar states with lifetimes of 10 ns to 100 ns as seen in the cluster time-series (Figure 11A, B). Cluster 1 in Figure 11B is comprised of ICL3 conformations very similar to the one found in the crystal structure. After about 200 ns (simulation #1) or earlier (#2, #3), no return of the ICL3 loop to its initial structure is observed, further corroborating that the ICL3 conformation in the respective crystal structures strongly depends on the contacts with the crystal lattice.

The RMSF plots show which residues are most flexible (Figure 11, bottom row) and how this changes from RhR to Ops* Figure 12. In RhR, the fluctuations suggest an approximate loop length of 8–10 residues with the residues around 236 fluctuating the most, which means the flexibility is shifted to the TM5 side. Altenbach et al. reported EPR measurements of individually spin-labeled amino acids within the ICL3 with the same results [162], the same shift in flexibility. Specifically, in Ops*, i. e. after activation, the TM5 side around residue K231 fluctuates less, whereas on the TM6 side the region around residue K248 shows increased flexibility. Finally, with bound $G\alpha_t$CT, the whole ICL3 region fluctuates only half as much as in Ops*, with RMSF values peaking at 4 Å instead of 8 Å.

3.1.2 *MD Simulation of RhR* and RhR*•$G\alpha_t$CT states*

At the start of the project, only the structures of the apoprotein Ops* and Ops*•$G\alpha_t$CT were determined, but not with covalently bound *all-trans* retinal. However, while we performed the simulations with the apoprotein, the structures of RhR* and RhR*•$G\alpha_t$CT were solved [66]. When parameters for the deprotonated *all-trans* retinal in RhR* [151]

became available, we took the opportunity and tried to verify the results obtained with Ops* and Ops*•Gα_tCT. And indeed, the TM6 tilts in simulations of RhR*•Gα_tCT (Figure A2D) exhibit distinctly narrower fluctuations than in simulations of RhR* alone (Figure A2F). This shows that in RhR*, too, the Gα_tCT has an ordering effect on the intracellular tip of TM6 and thereby on ICL3.

3.2 DYNAMICS OF TM6 IN β_2AR* AND RHR*

The β_2-adrenoceptor can activate multiple, distinct G proteins, coupling to both G_s and G_i [4]. While there are known ligands which can bias the receptors to couple to either [163], the structural mechanism by which the active β_2AR* can distinguish between G_s or G_i is unknown. Here we thus investigate the dynamics of TM6 as the source of the β_2AR* feature to bind to (and distinguish between) different G proteins. For that, we start by comparing available crystal structures of active GPCRs to see what structural differences there are and what might be the reason for them. We identify the TM6 tilt as the main difference and the bulkiness/slimness of the interaction partner as the likely cause. In MD simulations of uncomplexed β_2AR*, we observe different TM6 conformations in addition to a very flexible ICL3. This is in contrast to Ops* or RhR*, where there is considerable ICL3 flexibility but no TM6 conformation with a larger outward tilt (see Section 3.1, Figure A2F). Observing these different TM6 conformations (of which one looks very similar to the TM6 conformation of the crystal structures of RhR*•Gα_tCT) in β_2AR* raises the question if the different G proteins β_2AR is capable to bind can stabilize distinct TM6 conformations.

Peptides as surrogates

However, before performing MD simulations that directly address the question of TM6 conformations, we establish the use of peptides in MD simulations as a surrogate for the G protein holocomplex. The rationale behind this is twofold. First, it allows comparison with the available structural data of RhR*•Gα_tCT. Secondly, using peptides instead of the G protein holocomplex immensely lowers the computational load and allows performing computational analyses which otherwise would take prohibitive amounts of computing time. Our data suggests that peptides of sufficient length adequately mimic the coupling interface on the G protein side.

3.2.1 *Binding modes and space requirements of GαCT*

The crystal structures of β_2AR*•G$\alpha_s\beta\gamma$ [32] and of RhR*•Gα_tCT [31, 66, 67] are the only complexes reported so far that show the interaction of the GαCT and the R* binding crevice. Although the RhR*•Gα_tCT complex does not contain the G holoprotein, the structure and interaction of Gα_tCT agrees remarkably well with the Gα_sCT in β_2AR*•G$\alpha_s\beta\gamma$. Moreover, GαCT is a key determinant of R* G coupling specificity, which in case of rhodopsin and Gα_tCT stabilizes the same active MIIbH$^+$ receptor conformation as G_t [100], corroborating the use of GαCT peptides as surrogates of the holoprotein. Both Gα C-termini adopt an α-helical conformation terminated by a reverse

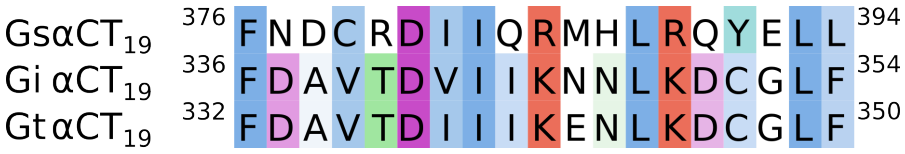

Figure 13: Alignment of GαCT19 sequences colored by similarity and residue type using the `clustelx` color scheme. Alignment columns are colored according to the majority consensus discerning hydrophobic (blue), basic (red), acidic (magenta), polar (green) and small (white) residue types. A non-unanimous consensus is colored lighter with non fitting outliers in white. Tyrosine is always colored turquoise. *The figure and legend are adapted from [101]*

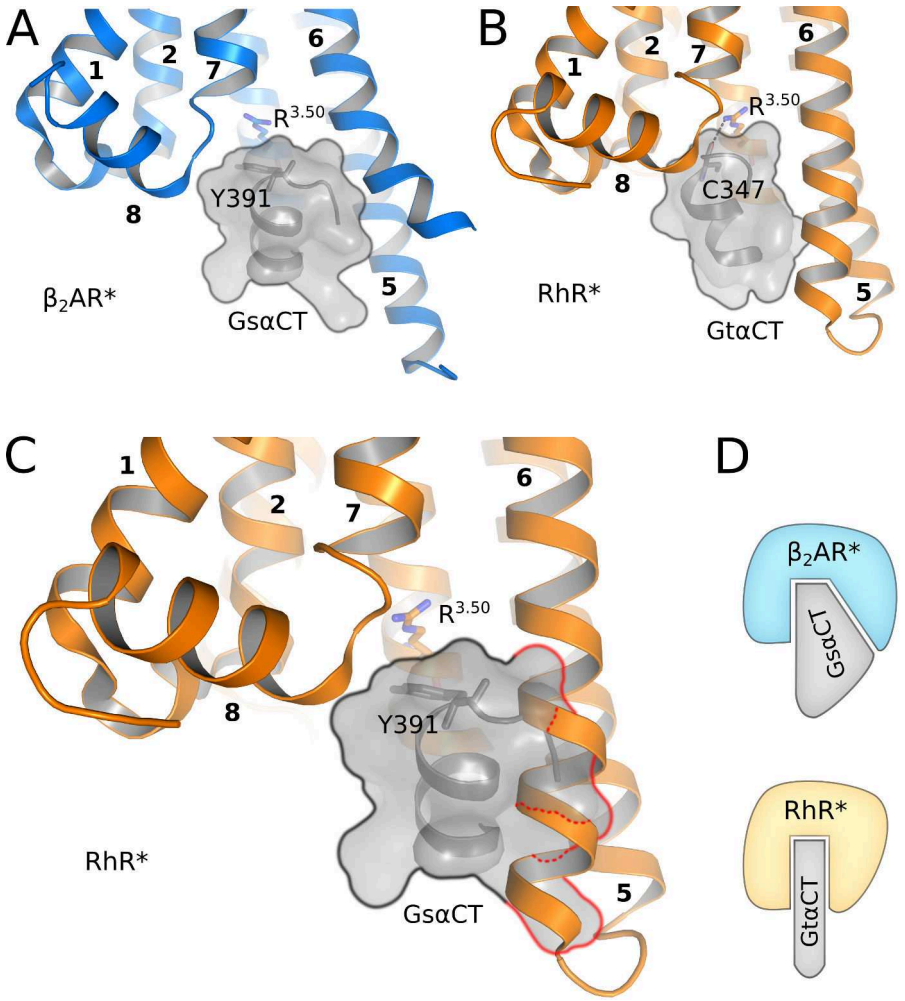

Figure 14: Cytoplasmic crevice of β_2AR* (blue) and RhR* (orange) from complexes with G_s [32] and $G\alpha_t$CT [31], respectively. (A-C) Side view: The outward-tilted TM6 of β_2AR* allocates space for the bulky $G\alpha_s$CT (outlined, translucent surface) that would clash (indicated by the red outline in C) with the more inward-tilted TM6 position of RhR* (orange) in complex with $G\alpha_t$CT (translucent surface). $R^{3.50}$ from the base of the cytoplasmic crevice and Y391/C347 from the tip of $G\alpha_{s/t}$CT are shown as sticks. (D) Schematic representations of the two crystal structure interactions in A-C. *The figure and legends are reproduced from [101]*

turn, whose tip binds to $R^{3.50}$ from the E(D)RY motif that forms the base of R* binding crevice. A sequence alignment shows bulkier side chains in the reverse turn of $G\alpha_sCT$ than in $G\alpha_tCT$ or its close homologue $G\alpha_iCT$ (Figure 13). Specifically, these are Y391 in $G\alpha_sCT$ vs R135 in $G\alpha_tCT$ or $G\alpha_iCT$. The $\beta_2AR^*{\bullet}G\alpha_s\beta\gamma$ complex structure then shows a cation-π interaction between Y391 and $R131^{3.50}$ (Figure 14A). Whereas, in RhR*•$G\alpha_tCT$, a hydrogen bond between the carbonyl oxygen of C347 and the guanidinium group of $R135^{3.50}$ connects the $G\alpha$ reverse turn with receptor-binding crevice (Figure 14B).

In addition to the cation-π interaction, the $\beta_2AR^*{\bullet}G\alpha_s\beta\gamma$ complex features a 5 Å to 6 Å larger outward tilt of TM6 as RhR*•$G\alpha_tCT$. Superposition of $G\alpha_sCT$ onto the position of $G\alpha_tCT$ in RhR* results in clashes between $G\alpha_sCT$ and TM6, suggesting that $G\alpha_sCT$ requires a larger outward tilt than $G\alpha_tCT$ or $G\alpha_iCT$ (Figure 14C). We thus propose that different space requirements for the key interactions of $G\alpha_tCT$ and $G\alpha_sCT$ with R* determine distinct TM6 outward tilts in the corresponding complexes (Figure 14D).

3.2.2 *Flexible TM6 behavior in uncomplexed β_2AR*

The crystal structure of the $\beta_2AR^*{\bullet}G\alpha_s\beta\gamma$ complex shows a TM6 conformation of β_2AR* tilted further outward than in RhR*•$G\alpha_tCT$. To find out whether β_2AR* can exists in a RhR*-like conformation, we removed the constraining $G\alpha_s\beta\gamma$. We then hypothesized that $G\alpha_iCT$ might be able to stabilize the TM6 of β_2AR* in a conformation similar to that in RhR*•$G\alpha_tCT$. Using the receptor coordinates from the $\beta_2AR^*{\bullet}G_s$ complex [32], we started a series of MD simulations of the uncomplexed β_2AR*. For that we reconstituted β_2AR* *in-silico* into a DMPC lipid bilayer solvated in water and carried out five independent 200 ns to 400 ns MD simulations.

With the cytoplasmic binding partner absent, we observe a broad distribution of TM6 tilts (Figure 15A,D), which is in accordance with earlier experimental [55] and computational analyses [54]. The TM2 to TM6 distance ($d_{TM2–6}$, see Section 2.4.4) peaking at 28 Å reflects the presence of the G_s-coupled state in the uncomplexed β_2AR*. However, at around 23 Å, another peak arises, defining another population of TM6 tilts in β_2AR* (Figure A1A). The smaller $d_{TM2–6}$ in that population renders it similar to the more closed RhR*•$G\alpha_tCT$ state. During the simulations, we observe TM6 moving back and forth between the states of both populations, while the cytoplasmic end of TM6 undergoes pronounced order-to-disorder transitions (Figure A1B). Such flexibility of R* has been previously shown by spectroscopic studies applying NMR on β_2AR* [164] or FTIR [43] on RhR*.

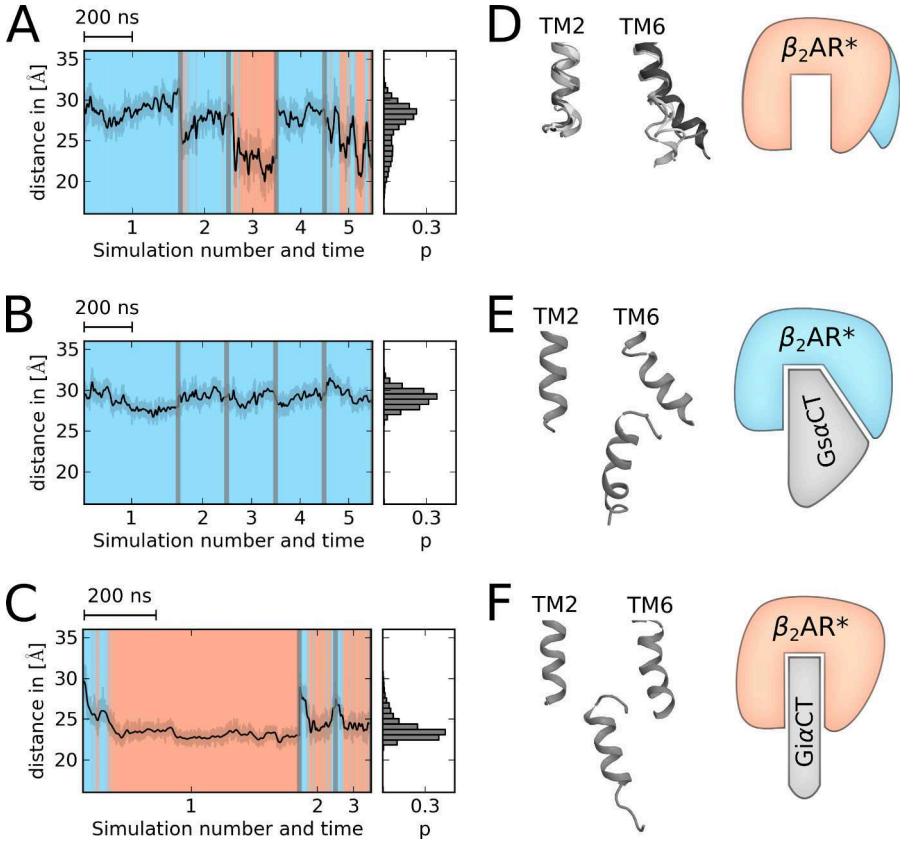

Figure 15: Conformations of the cytoplasmic crevice of β_2AR* observed in MD simulations (A,D) alone, (B,E) with Gα_sCT19, and (C,F) with Gα_iCT19. Panels A-C show the time traces of the TM6 tilts ($d_{TM2–6}$) with resulting distributions and observed probabilities p. Panels D-F schematically depict the Gs-coupled state (blue) and the RhR*-like state of β_2AR* (red) defined by representative TM2 and TM6 conformations. (C) In 3 of 21 simulations (100 ns to 600 ns), β_2AR* adopts a RhR*-like conformation with Gα_iCT19. Crystal structures of R and R* exhibit the following TM6 tilts: inactive β_2AR (PDB entry 2RH1), 18.2 Å; β_2AR*•G_s (3SN6), 29.7 Å; inactive RhR (1U19), 16.4 Å; and RhR*•Gα_tCT (3PQR), 23.2 Å. *The figure and legend are reproduced from [101]*

3.2.3 *Using GaCT peptides as a surrogate of G*

GαCT peptides have been successfully used as surrogates for G to stabilize the active receptor conformation [76]. Recently, G$\alpha_{s/i}$CT peptide based FRET sensors have been used to detect a β_2AR* ligand that selectively activate G_i over G_s [163]. The β_2AR*•G$\alpha_s\beta\gamma$ complex gives us the opportunity to identify the minimum length a Gα_sCT peptide needs to sufficiently mimic complex stabilization by G_s so that it can be used as a surrogate of G_s. For that, we performed a series of β_2AR* MD simulations with G_s truncated to Gα_sCT 19-mer (Gα_sCT19, ^{376}FNDCRDIIQRMHLRQYELL^{394}) and 11-mer peptides. The effect of the peptides on TM6 in these simulations was then compared to simulations of the full β_2AR*•G$\alpha_s\beta\gamma$ complex.

The 19-mer peptide remains in its starting conformation during all simulations (Figure A3) and keeps TM6 at its initial position (Figure 15B,E) through the same specific interactions as with G_s (Figures A4B and A5B). However, with the further truncated 11-mer peptide, we observe a higher positional variability of $G\alpha_s$CT11 (Figure A3) and TM6 (Figure A2) stemming from a reduction of specific contacts between β_2AR* and $G\alpha_s$CT11 (Figures A4C and A5C).

The higher positional variability of $G\alpha_s$CT 11 compared to $G\alpha_s$CT 19 and an analysis of the specific contact between $G\alpha_s$CT and β_2AR* suggest that 15 C-terminal residues of $G\alpha$ are required to form the binding interface for G_s coupling to β_2AR* (see Figures 17A, A4A-C, and A5A-C). These 15 residues can form a scaffold between ICL2 and TM5/6 that alone sufficiently stabilizes the more open cytoplasmic crevice of β_2AR* coupled to G_s. But the initially helical N-terminus of the $G\alpha_s$CT11 peptide exhibits some unwinding, too. To account for that unwinding, we will use GαCT19 peptides, not 15-mer peptides, as surrogates of G to investigate the coupling to β_2AR* and in particular the stabilization of specific TM6 conformations.

3.2.4 *Stabilization of a RhR*-like conformation of β_2AR* by $G\alpha_i$CT*

Given the observation of a RhR*-like conformation in the ensemble of uncomplexed β_2AR* conformations, we addressed the question of whether the RhR*-like conformation can be stabilized by a $G\alpha_i$ C-terminal 19-mer ($G\alpha_i$CT19). Finding that β_2AR* can accommodate $G\alpha_i$CT with a RhR*-like conformation would corroborate the proposed role of TM6 in distinguishing between $G\alpha_s$CT and $G\alpha_i$CT. To address this question, we extrapolated a starting position for $G\alpha_i$CT19 (^{336}FDAVTD*V*IIKNNLKDCGLF^{354}) from the crystal structure complex of $G\alpha_t$CT19 (^{332}FDAVTD*I*IIKENLKDCGLF^{350}) with RhR* [31] (see Section 2.2.2). The receptor coordinates were taken from the β_2AR*•G_s crystal structure complex [32] to form a β_2AR*•$G\alpha_i$CT complex suitable for starting MD simulations where $G\alpha_i$CT19 initially does not have any contact with TM6. In 18 of the 21 independent simulations started with that complex, TM6 tends to persist in its initial position for the first hundreds of nanoseconds (Figure A2E). This TM6 behavior is similar to what we observed in the simulations of uncomplexed β_2AR* (Figure 15A, D). However, the initial receptor conformation is left and TM6 spontaneously tilts inward by 6 Å (Figure 15C,F) in 3 simulations. In those simulations, a β_2AR*•$G\alpha_i$CT complex with a $d_{TM2–6}$ = 23 Å is formed.

β_2AR•$G\alpha_i$CT with a G_s-like TM6 conformation*

The 18 remaining simulations exhibit no departure from the initial, crystal structure conformation of the receptor. There TM6 stays in an outward-tilted conformation, but the $G\alpha_i$CT peptide changes. Leav-

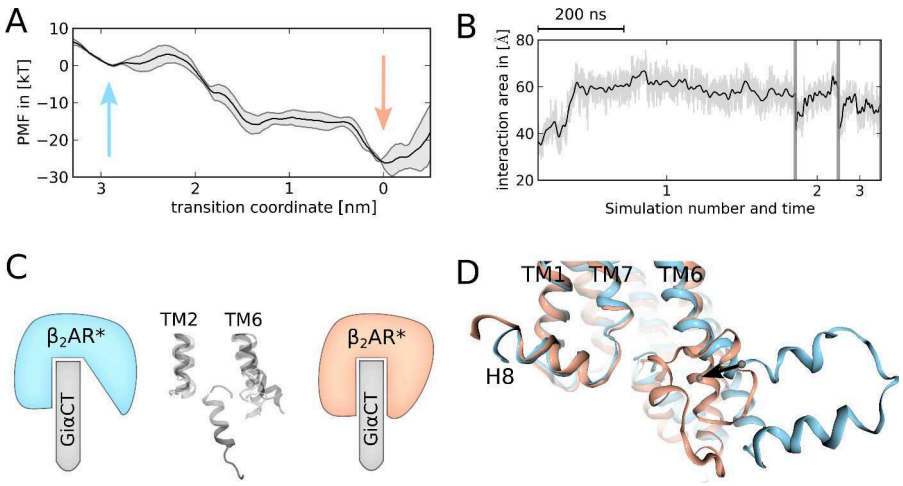

Figure 16: Umbrella sampling run along the coordinates of the TM6 inward movement during transition from the starting state using the receptor coordinates from the β_2AR*•G_s complex to the β_2AR*•Gα_iCT 19 complex (see Figure 15). (A) potential of mean force (PMF) energy profile (black line) and error estimation (gray area). Arrows indicate the positions of the initial (G_s coupling state, blue) and the final β_2AR* conformation (red) on the transition coordinate (Section 2.3.3). An energy barrier of (3.0 ± 2.3) kT is at 2.3 nm on the transition coordinate, which corresponds to $d_{TM2\text{–}6} \approx 28.6$ Å. (B) Increase of the β_2AR* Gα_iCT19 interaction surface accompanying the TM6 inward movement. (C) Schematic depiction of the TM6 inward movement and corresponding (D) view of the the intracellular binding crevice, using the actual receptor coordinates from the simulation. *Panel (A), (B) and their legend are reproduced from [101]*

ing its initial position, Gα_iCT gains contact with TM6 by adopting a Gα_sCT-like conformation. After Gα_iCT is wedged that way in the binding crevice, we unsurprisingly do not observe TM6 tilting inward as it is now stabilized by the peptide. Presumably, this conformation is heavily biased by the initial crystal structure conformations, of the receptor from which the simulation starts. Nevertheless, Gα_iCT seems to be able to bind in two different conformation depending on the TM6 position. While this raises the question why there may be two different Gα_iCT binding modes, it does not affect the observation that Gα_sCT exhibits only one binding mode.

Umbrella sampling

Observing the TM6 inward motion in just a fraction of the performed simulation, points to an substantial energy barrier that retards formation of the more closed β_2AR* state that binds Gα_iCT. To characterize that energy barrier, we performed umbrella sampling (US) MD simulations along one of the trajectories that lead to a β_2AR*•Gα_iCT complex. For this trajectory, we calculated the PMF (Figure 16A). The PMFs firstly confirm the existence of such an energy barrier which probably arises from the reorganization of interactions that is ob-

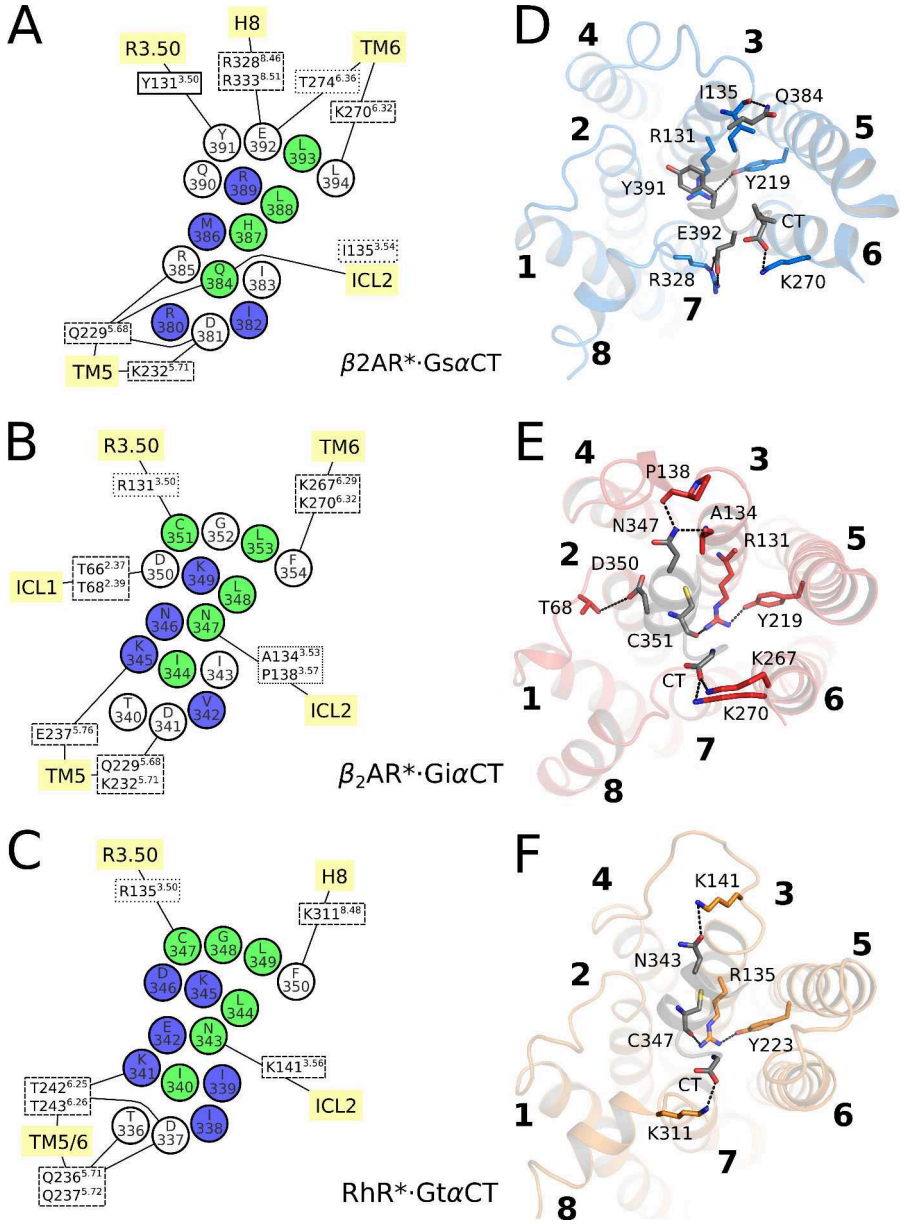

Figure 17: Contacts observed in MD simulations of various R*•GαCT19 complexes. (A-C) Recurring polar contacts involving side chains (dashed boxes), backbone carbonyl (dotted boxes), or cation-π interactions (solid boxes) between peptides (circles) and receptors (boxes) depicted as interaction schemes. Buried residues are colored green; accessible residues are blue (see also Figure A8). (D-E) Top view from the intracellular side on the different R*•GαCT19 complexes. Hydrogen bonds between residues (sticks) of peptides (gray) and R* (colored) are indicated as dotted lines. For clarity, only the C-terminal 11 residues of the GαCT19 peptides are shown in panels D-F, but in panels A-C, the full set of interactions is described. *The figure and legend are reproduced from [101]*

served between TM5/6 (Figure A6B). Secondly, the PMFs show a continuous downhill reaction for the TM6 inward movement after the barrier is overcome.

While $G\alpha_i$CT19 remains at its starting position and forms the key interaction between C351 and $R131^{3.50}$, the interaction surface between β_2AR* and $G\alpha_i$CT19 increases (Figure 16B). Additional specific interaction in form of hydrogen bonds between N347 and $P138^{3.57}$ from the second intracellular loop (ICL2) and from D341 and K345 to ICL3 are also observed. Notably, these interactions reflect the hydrogen bond network that is observed in MD simulations of $G\alpha_t$CT with RhR* (Figures 17C and A5E). As a result, β_2AR* can form a complex with $G\alpha_i$CT that shows remarkable similarity to the crystallographically determined RhR*•$G\alpha_t$CT complex (Figure A7C). The TM6 is tilted inwards, closing the cytoplasmic crevice and forming a tight interaction with $G\alpha_i$CT (Figure 16B). The resulting complex of the extrapolation of $G\alpha_i$CT to G_i (based on β_2AR*•G_s) has no clashes with the membrane nor within the complex itself (Figure A7).

Comparison of R•GαCT19 complexes*

Including the RhR*-like β_2AR*•$G\alpha_i$CT complex in addition to the β_2AR*•$G\alpha_s$CT and the RhR*•$G\alpha_t$CT complex, we have now obtained simulation data on three distinct receptor GαCT complexes. To describe the observed differences of their respective coupling interface, we analyzed the hydrogen bonding pattern, cation-π interactions (Figure A5) and the interaction area (Figure A8) between receptor and peptide. The results are summarized in Figure 17. Apart from the already described interaction to $R^{3.50}$, all complexes show multiple hydrogen bonds from the peptide TM5/6 and ICL2. We find hydrogen bonds to H8 in β_2AR*•$G\alpha_s$CT and RhR*•$G\alpha_t$CT but not in β_2AR*•$G\alpha_i$CT, which in turn exhibits hydrogen bonds to ICL1 exclusively. With 4 buried and 4 exposed residues, the $G\alpha_s$CT peptide is less enclosed in our simulations as $G\alpha_i$CT (5 buried, 4 exposed) and $G\alpha_t$CT (6 buried, 6 exposed). This clear-cut distinction between buried and exposed residues in RhR*•$G\alpha_t$CT suggest a very specific coupling interface.

3.2.5 *Control MD simulations of RhR* with and without $G\alpha_s$CT*

Having established that β_2AR* is capable of accommodating both $G\alpha_s$CT and $G\alpha_i$CT peptides, we wondered whether RhR* can do it too, that is form a well-defined complex with $G\alpha_s$CT in addition to RhR*•$G\alpha_t$CT for which crystal structures exist. While it has been shown that RhR* does not form functional complexes with G_s [165], investigating a putative binding mode of $G\alpha_s$CT with RhR* is still of interest. In the beginning of this section, we propose that the dynamics of TM6 are the crucial feature of β_2AR* for distinguishing between different G proteins. Finding that RhR* lacks similar TM6 dynamics and fails to accommodate $G\alpha_s$CT would serve as a control of this hypothesis, corroborating the role of TM6. To this end we

performed a series of MD simulations with RhR* and (modeled) putative RhR*•Gα_sCT complexes. In this RhR*•Gα_sCT model, Gα_sCT is bound in a way resembling Gα_sCT bound to β_2AR* in the β_2AR*•G$\alpha_s\beta\gamma$ complex (Section 2.2.2).

First, we do not observe larger TM6 outward tilts corresponding to the β_2AR*•G_s conformation in MD simulations of uncomplexed RhR* (Figure A2F). But for β_2AR* we did observe RhR*•Gα_tCT-like conformations of TM6. Secondly, in simulations of a putative RhR*•Gα_sCT peptide complex (see Section 2.2.2), the characteristic cation-π interaction between Y391 and $R^{3.50}$ (here R135) does not form as with β_2AR*•Gα_sCT (A4). Taken together, these results suggests that TM6 tilts in RhR* are limited and that RhR* can not accommodate Gα_sCT. This is in accordance with RhR*'s failure to activate G_s [165] and the absence of a G_s signaling pathway in the visual rhodopsin system [12].

3.3 ROLE OF R* IN G α5 DISPLACEMENT

The intermediary R*•G[GDP] is the first stable/specific complex of receptor and G protein in the sequence of events during receptor-triggered G protein activation (Figure 5). R*•G[GDP] accumulates under excess of GDP while GTP is absent [44, 83], likely involving the α5 helix for coupling [83]. Here we investigate how the transition from the intermediary R*•G[GDP] to the nucleotide-free R*•G[empty] complex causes nucleotide release. Specifically, what the role of the α5 helix at the receptor G protein coupling interface is. We propose that the α5 helix rotates at/within the coupling interface during the transition leading to the nucleotide release as described in [78].

Due to the lack of structural data on the intermediary R*•G[GDP], we first create a model of that complex based on flexible docking calculations and requirements deduced from experimental data on the R*•G[GDP] and R*•G[empty] complexes and their transition. Starting from the intermediary R*•G[GDP] model, we employ MD simulations to test if the α5 helix alone (without the rest of the G protein) can perform its proposed role at the receptor G protein coupling interface.

3.3.1 *Prerequisites of a R*•G[GDP] complex*

The intermediary R*•G[GDP] complex should meet a number of requirements deduced from experimental data and sterical constraints placed on the G orientation by the membrane environment, the preceding and the following G conformations.

The G protein in the uncomplexed Gα[GDP] and in the R*•G complex differ in a number of places as revealed by X-ray structure and HDX as well as EPR analyses [80, 32, 81, 82, 83]. The α5 helix switches (i. e. rotates and translates with respect to the Gα Ras domain) and the all-helical (AH) domain is displaced. Directly contacting the GDP nucleotide and forming its binding pocket, the loop at the N-terminal end of the α5 helix and the AH domain are crucial for GDP stabilization. To maintain a stable GDP binding pocket it is thus reasonable that in the nucleotide-bound R*•G[GDP] intermediate the structure known from uncomplexed Gα[GDP] is preserved, i. e. the α5 helix is not switched (and the all-helical domain is not displaced).

For the β_2AR* Gα_s system, HDX experiments suggest that Gα_s with GDP bound couples to β_2AR* mainly through contact with Gα_sCT [83]. After adding GDP to a nucleotide-free preparation that stabilized the empty site complex, the exchange rates do only slightly increase at the α5 C-terminus. These exchange rates show differences between GDP-bound and nucleotide-free complexes. Low exchange rates indicate structured elements, which suggests that α-helical con-

formation of $G\alpha_s$CT in GDP-bound $G\alpha_s$ that is stabilized through interactions with β_2AR*. In contrast, adding a non-hydrolyzable GTP analog uncouples G_s from β_2AR*•G_s and results in high exchange rates of the α5 C-terminus [83].

Superposing the crystal structures of uncomplexed G, $G\alpha_s$[GTPγS] or $G\alpha_i$[GDP] with R*•GαCT shows major clashes with the membrane (see Fig. 5 of ref. [31]). Meeting the requirement of a not switched α5 helix to preserve the structure of uncomplexed Gα[GDP] and maintain a stable GDP binding pocket, the C-terminus of GDP bound Gα must thus bind with a different orientation to β_2AR* in the GDP-bound as in the empty site complex. Of note, rotating the complete G protein by the amount the α5 helix turns in the helix switch, a clash free model of the R*•G_s[GDP] complex can be obtained as shown previously for RhR*•G_t [78].

3.3.2 *R*•G[GDP] complexes identified by flexible docking and modeling*

With the aim of modeling R*•G[GDP] complexes, docking calculations were performed to detect feasible $G\alpha_s$CT interactions with the β_2AR* cytoplasmic crevice. To this end, P. Hildebrand applied the flexible docking protocol as previously described [78], using a fixed α-helical backbone geometry but flexible side chains. The thus obtained docking poses were then analyzed by clustering as described in Section 2.5.

Previously, docking of $G\alpha_t$CT 15- and 19-mer peptides recovered the position of the RhR*•$G\alpha_t$CT X-ray structure, which is likely the position α5 would adopt in the RhR*•G_t empty site complex. In addition, a second position featuring different interactions was found and assigned to reflect the intermediary RhR*•G_t[GDP] complex [78]. Now, in a flexible docking analysis of 15-mer $G\alpha_s$CT to β_2AR*, we find an analogous pair of states for the β_2AR* G_s system (Figures 19, 18).

$G\alpha_s$CT docking states

The flexible docking yields multiple states, of which the highest scoring conforms to the position and orientation of $G\alpha_s$CT seen in the co-crystal of β_2AR*•G_s (Figure 18), thus confirming the applicability of the docking approach. Moreover, this state shows the characteristic cation-π interaction between Y391 and R131$^{3.50}$ (Figure 19B). A number of contacts from the N-terminus of 15- mer $G\alpha_s$CT to ICL2 and 3 of β_2AR* are also formed as in the β_2AR*•G_s complex. Notably, hydrogen bonds of Q384 with the main chain carbonyl group of I135 from the conserved P138$^{3.57}$ C-terminal cap of TM3 (Figure 19D) are observed. A second docking state of $G\alpha_s$CT has fitting properties for a β_2AR*•G_s[GDP] model. There, a rotation around the helix axis (i. e. compared to the first docking state) results in a shift of the reverse

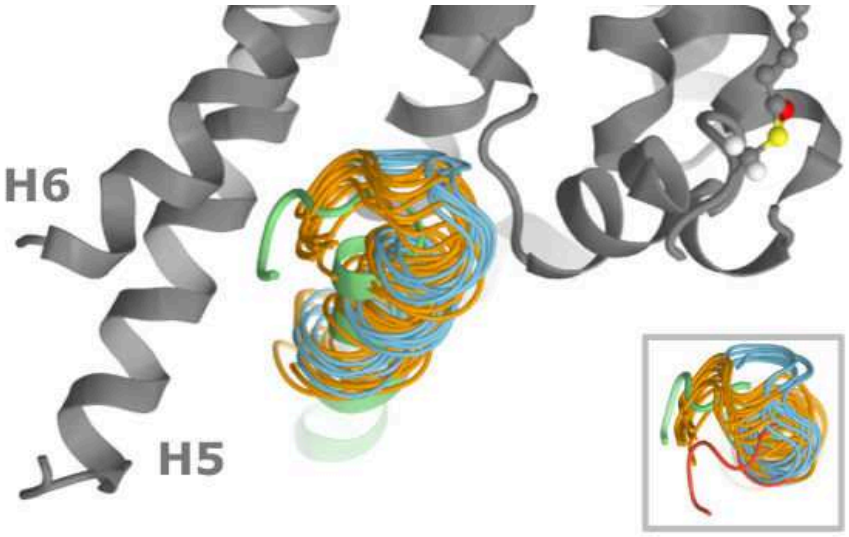

Figure 18: Clusters of peptide positions obtained by flexible docking of 15-mer $G\alpha_s CT$ to $\beta_2 AR^*$. The largest cluster (orange) coincides with the position and orientation of $G\alpha_s CT$ in the X-ray structure. However, in agreement with the MD simulations of $\beta_2 AR^* \bullet G\alpha_s CT$ (A10) $G\alpha_s CT$ is shifted away from TM3 as compared to its position in the $\beta_2 AR^* \bullet G_s$[empty] complex. In the blue cluster (third largest cluster), $G\alpha_s CT$ is rotated by 70° compared to the X-ray state and represents the position in a putative R*• G[GDP] intermediate. The inset shows the view along the helix axis, highlighting the difference between both states. The orientation of $G\alpha_s CT$ from the second largest cluster (representative structure shown in red in the inset) is neither compatible with the X-ray state nor with the R*•G[GDP] intermediate because it over-twists the rotation of $G\alpha_s CT$. Hypothetically, this position refers to one of the intermediate dissociation states observed by single particle EM of the $\beta_2 AR^* \bullet G_s$ complex [82].

turn of $G\alpha_s CT$ by one residue relative to $R131^{3.50}$. The cation-π interaction with Y391 is replaced by a hydrogen bond $R131^{3.50}$ forms with E392 (Figure 19A, B). In lieu of a hydrogen bond to Q384 at the N-terminus, a potential hydrogen bond between the backbone oxygen of I135 from the $P138^{3.57}$ cap and R385 forms (Figure 19C, D).

In the described docking states, the same conserved structural motifs of $\beta_2 AR^*$ are employed by $G\alpha_s CT$ for binding. However, comparison of both states shows that the shift of the interacting residues of $G\alpha_s CT$ by one position goes along with a clockwise rotation of $G\alpha_s CT$ by 60° and a translation of 1.5 Å. Given that the α5 helix within G_s is rotated in the opposite direction by the same magnitude between the uncomplexed Gα[GDP] and the R*•G complex, the nucleotide-bound states of G_s or G_i can be superposed with $G\alpha_s CT$ of the second docking state to create a very similar complex configuration as in the $\beta_2 AR^* \bullet G_s$ complex, however, with α5 rotated by 60° (Figure 19E, F). As in the previous analysis of the RhR* G_t interaction [78], the such created complex does not cause any major protein-

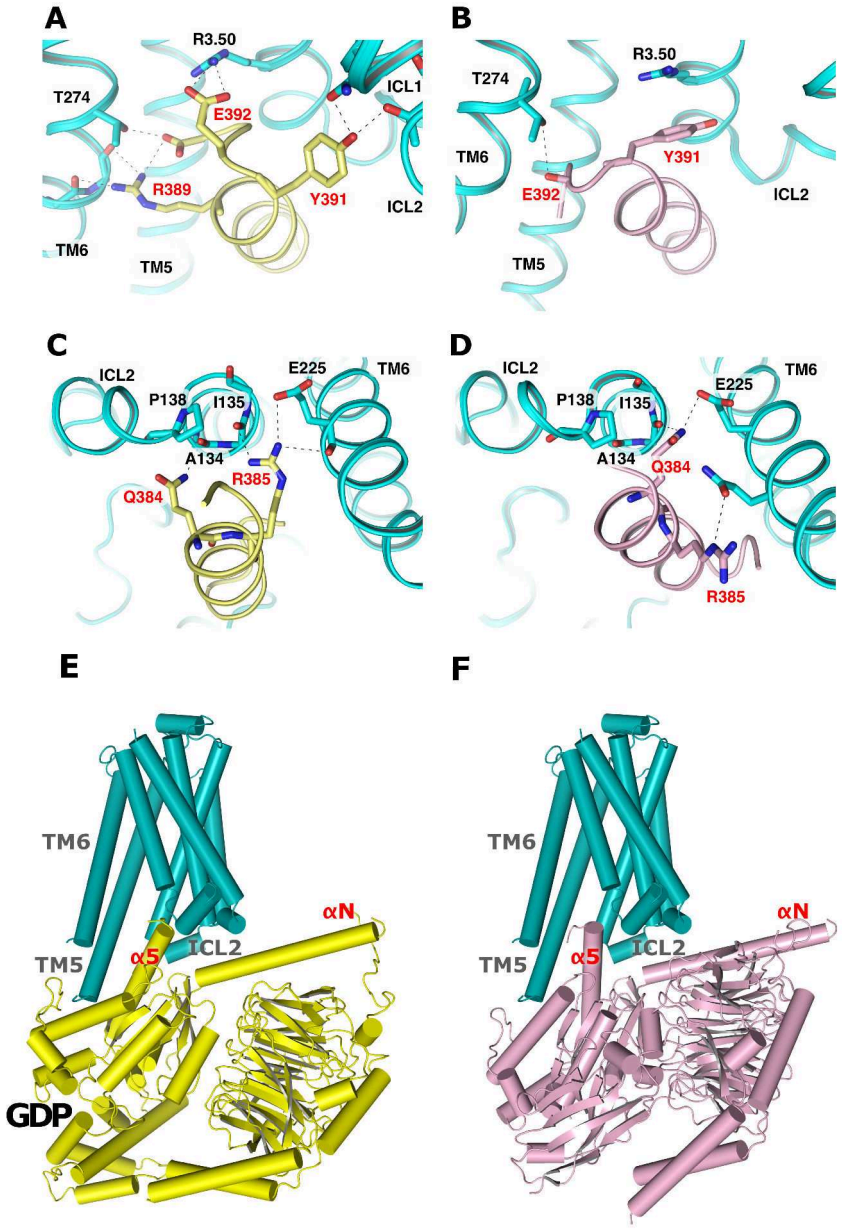

Figure 19: Comparison of (left panel) the β_2AR*•G_s[GDP] model and (right panel) the β_2AR*•G_s X-ray structure. Potential hydrogen bonds to the cytoplasmic crevice (cyan cartoon) from (A, B) the C-terminal reverse turn and (C, D) the N-terminus of Gα_sCT in (A, C) the intermediate position obtained from flexible docking (yellow cartoon) and (B, D) in the empty site complex (magenta cartoon), respectively. Residue labels from β_2AR* are colored in black, residues from Gα_sCT are red. Potential hydrogen bonds are denoted as black dashed lines. (E) Complete model of the β_2AR*•G_s[GDP] intermediate compared to (F) the β_2AR*•G_s X-ray structure (PDB entry 3SN6). R*•G[GDP] was obtained by superposition of Gα_s[GTPγS] (PDB entry 1AZT), with the intermediary β_2AR*•Gα_sCT complex obtained from flexible docking analysis by common backbone atoms. $G_s\beta$- and $G_s\gamma$ of PDB entry 3SN6 were superimposed to Gα_s to obtain a complete model. *The figure and legend are reproduced from [166]*

protein/protein-membrane clashes or distortions. Since in that complex α5 is also structured and the G_s is not altered, especially at the the GDP binding pocket, it fulfills the above-listed requirements and was assigned to the β_2AR*•G_s[GDP] complex.

3.3.3 *Observation of GαCT switches at the R* G protein interface*

The rotation of α5 within the cytoplasmic crevice of the R*•G[GDP] complex is the distinguishing feature that separates the GDP-bound from the nucleotide-free state. With the R*•G[GDP] complex at hand, we now want to test whether the rotation of α5 can result from interactions of GαCT and the cytoplasmic crevice alone, that is without requiring the G holoprotein to be present. To perform the test, we started 30 MD simulations from the R*•GαCT complexes from the intermediary GαCT positions obtained by flexible docking of both systems — the β_2AR* $G\alpha_s$CT as well as the RhR* $G\alpha_t$CT system. Since we work with the peptides here and not the holoprotein, only the interactions between GαCT (i. e. α5) and R* can influence the test as desired. Additionally, we performed the simulations with 11-mer GαCT peptides which are the largest common structure of GαCT in the available experimental data [66, 167, 68, 69, 32, 31, 67]. However, as a test, we also evaluated the effect of the peptide length by performing 10 more MD simulations with 19-mer GαCT peptides. In general, all simulations were unbiased with GαCT neither restricted to its starting conformation nor in its mobility.

R•G[GDP] simulation with GαCT*

We performed a total of 60 MD simulations, 30 each for the β_2AR* and the RhR* systems, of which 18 feature a GαCT helix switch, 8 for $G\alpha_s$CT and 10 for $G\alpha_t$CT. Analysis of the MD simulations shows that the 11-mer GαCT adopts in about one third of the simulations a conformation maintained for the remainder of the simulation (Figures A22–A31). Formation of the maintained binding modes occurs quickly, i. e. within 50 ns for $G\alpha_s$CT and within 3 ns for $G\alpha_t$CT. With the exception of two simulations, we observe a switch-like transition of GαCT from its starting position to its position seen in the X-ray structures in all of them. The transitions are monitored by the parameters backbone RMSD and peptide rotation (Figures 20, A22, A23). Despite the fact that two different receptors and peptides were simulated, the GαCT peptides undergo similar screwing motions to reestablish the key interaction with $R^{3.50}$ as in the nucleotide-free state. The formation of that interaction becomes apparent in time-resolved analyses of distance and energy (Figures 20E-F, A24, A25). During the GαCT helix switch, the secondary structure of GαCT persists (Figures A26, A27) while a number of water molecules are displaced

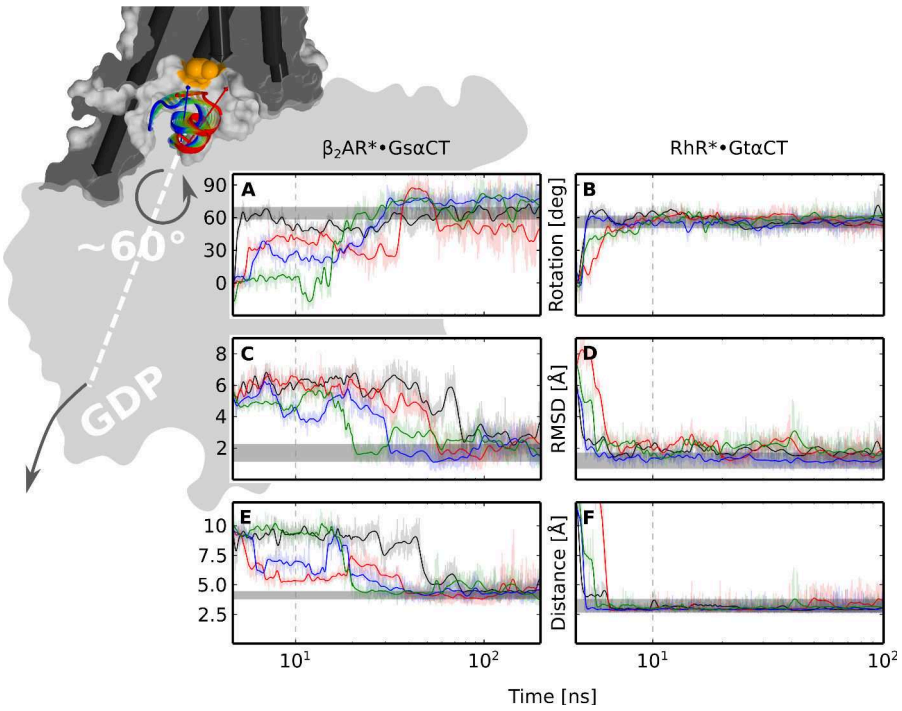

Figure 20: Switch of $G\alpha_s CT$ (left) and $G\alpha_t CT$ (right) at the R* interface observed in MD simulations. Background panel: $G\alpha_s CT$ switches with the cytoplasmic crevice of $\beta_2 AR^*$ from the intermediary (red) to the empty site position (blue). The transition is schematically indicated by semi-transparent colored cartoons. $G\alpha_s CT$ is rotated around its helix axis (red and blue arrows) by about 60°, which eventually leads to GDP release from the nucleotide binding pocket of the G_s holoprotein (gray, flat-shaded). In addition, a tilt motion of $G\alpha_s CT$ parallel to the membrane plane is observed. The surface of the receptor (gray) is cut at the position of $R^{3.50}$ (orange patch) located at the floor of the cytoplasmic crevice. TM helices are drawn as cylinders. For clarity, H8 and TM6 of $\beta_2 AR^*$ are omitted. Backbone-RMSD of (A) $G\alpha_s CT$ or (B) $G\alpha_t CT$ relative to their position in the X-ray structure. Change in rotation of (C) $G\alpha_s CT$ or (D) $G\alpha_t CT$ around its helix axis. Distance between (E) the center of the phenyl ring of Y391 of $G\alpha_s CT$ and $R131^{3.50}$ or (F) between the carbonyl oxygen of C347 of $G\alpha_t CT$ and $R135^{3.50}$. Gray bars indicate the mobility of GαCT in MD simulations of (left) $\beta_2 AR^*$•G_s or (right) RhR*•$G\alpha_t CT$ (see Figures A10, A11 and A13). The mobility of switched $G\alpha_s CT$ (after about 100 ns) is only slightly increased when compared to the mobility of the corresponding section in $\beta_2 AR^*$•G_s (gray bar). The time series data are drawn on top of the raw data as a running average. The plots are linear for the first 10 ns and logarithmic for the remaining time (gray dashed lines). The four representative simulations (black, red, blue, green) of $G\alpha_s CT$ (Figure A22, simulations 8, 9, 21 and 23) and of $G\alpha_t CT$ (Figure A23, simulations 9, 16, 21 and 30) were picked from 8 and 10 simulations where a switch was observed (Figures A22 and A23). *The figure and legend are reproduced from [166]*

from the coupling interface in favor of the formation of a hydrophobic patch between GαCT and ICL3 of R* (Figures A19, A28–A31).

In simulations where no single conformation is maintained, GαCT diffuses away and unfolds or fluctuates between different orientations and positions. The only additional maintained binding mode of GαCT is observed in simulations 11 and 12 in the RhR* $G\alpha_t$CT system (Figures A22B–A31B). There, the helix axis of $G\alpha_t$CT is tilted 10° steeper toward the membrane plane.

19-mer GαCT simulations

The additional simulations of 19-mer GαCT performed to the evaluate the effect of peptide length show very similar dynamics as observed with the shorter 11-mer peptides, including two helix switch events for each system respectively (Figure A18). Due to the lack of stabilizing contacts with R*, the N-terminal portion of 19-mer GαCT is highly mobile (Figure A18A). In contrast, the mobility of the C-terminal reverse turn of switched GαCT is low, as with the 11-mer peptides and comparable to the mobility of the corresponding section in β_2AR*•G_s (Figure 20A, C, E, gray bars).

4 DISCUSSION

4.1 DISORDERED REGIONS AND RHR*•Gα_tCT FORMATION

Active rhodopsin (RhR*) can activate up to 10^3 G proteins (G, G$\alpha\beta\gamma$) per second and more [44], while at the same time exhibiting virtually no constitutive activity. To elucidate the structural determinants that allow such a perfect switching behavior (i. e. unparalleled speed and fidelity), we investigated intrinsically unstructured regions in both binding partners. In the G protein the Gα C-terminus (GαCT) becomes structured upon binding to RhR* [68, 31, 32]. For the receptor, the third intracellular loop (ICL3) is a candidate for an intrinsically unstructured region, as ICL3 is often unstructured in other GPCRs and sequence analysis, too, suggests it is intrinsically unstructured [61]. However, structures of rhodopsin obtained by X-ray crystallography show the ICL3 of rhodopsin and its apoprotein opsin resolved in both inactive and active receptor states and regardless of being bound to a Gα_tCT peptide or not [39, 30, 31].

4.1.1 *ICL3 structure and effect of Gα_tCT observed in MD simulations*

The molecular dynamics (MD) simulations of rhodopsin show the fluctuations of the receptor structures in the first few hundred nanoseconds after removal from the crystal lattice and placement in a membrane-water environment (Figure 21). Within that timespan, the ICL3 of the receptor is generally observed to be flexible unless constrained by a Gα_tCT peptide. In simulations of RhR, the flexibility reflects the high B-factors seen in the respective crystal structures. In contrast, the simulations of RhR* without a Gα_tCT peptide show high ICL3 flexibility despite moderate B-factors in the respective crystal structures. However, in the crystal structures of RhR*, the ICL3 is stabilized by a number of contacts within the crystal lattice: the symmetry mates contact ICL3 with *ICL2* and *helix 8* (*H8*, Figure 21). The lack of those constraints (resulting from the experimental conditions) in the MD simulations fits to the differing ICL3 behavior in the simulations.

Interpretation of the observations from the MD simulations should be done with respect to the ensemble of ICL3 conformations at statistically robust timescales. For that, we compare the MD data with Fourier transform infrared spectroscopy (FTIR) results on the conformational diversity of RhR*. The FTIR experiments provide information on the ensemble of ICL3 conformations for the receptor states

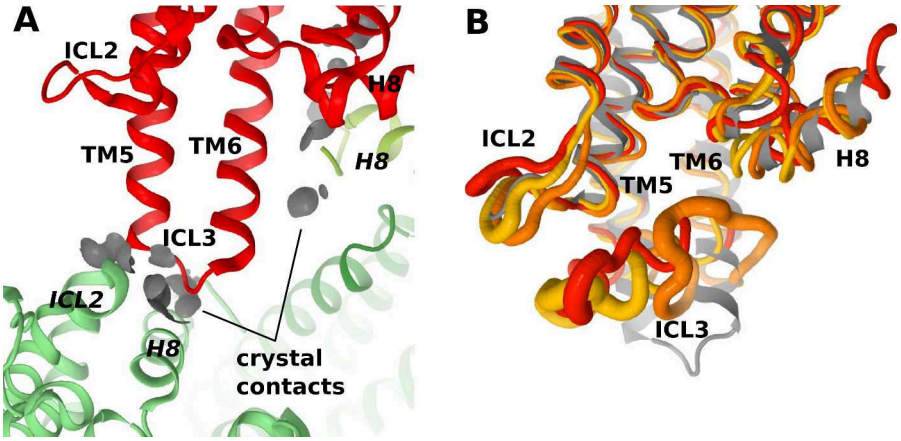

Figure 21: Comparison of RhR* ICL3 crystal lattice contacts and flexibility observed in MD simulations. (A) Crystal contacts in RhR* (PDB entry 3PXO) involving ICL3 are shown as gray patches between a central receptor molecule (red) and its symmetry mates (green). The symmetry mates contact ICL3 with *ICL2* and *H8* as well as TM6 with *H8* (italic labels). (B) Flexibility of ICL3 observed in MD simulations of RhR* embedded in a lipid bilayer-water environment. Shown are the crystal structure (gray), which is also the starting conformation, and three representative conformations (red, orange, yellow) obtained by cluster analysis.

that populate each of the active receptor's conformations. However, assigning the distinctive FITR markers to structural changes is not readily possible, especially at atomic resolution. By combining the FTIR data on well-defined molecular ensembles with the atomistic MD data of single molecules, a more complete view of the ICL3 behavior in RhR* is provided [56].

4.1.2 *FTIR spectroscopy links stabilizing effect of $G\alpha_t$CT to ICL3*

Using FTIR spectroscopy, the metarhodopsin states of light-induced retinal-bound receptors can be distinguished (for an in depth investigation of rhodopsin with FTIR see Elgeti [168]). Within milliseconds of illumination, rhodopsin leaves the dark state conformation and enters the multi-step equilibrium scheme of metarhodopsin states (Scheme 3). The exact distribution of metarhodopsin states across the equilibrium can be controlled by experimental conditions, i. e. by pH and temperature. By analyzing pH titration curves of specific FTIR difference bands using Scheme 3, the fractions of the individual metarhodopsin states can be deduced [169, 56]. At high pH (> 9) and a temperature of 30 °C, MI, MIIa, and MIIb are found to be each populating a fraction of approximately one-third. However, at low pH, the titration curves converge with an apparent pK_a of 7.5, indicating stabilization of the protonated $MIIbH^+$ receptor state. In the pH range of 5 to 6, the $MIIbH^+$ receptor state is selectively stabilized by

facilitating proton uptake to E134 and thus destabilizes the inactive state [170].

$$\text{MI} \rightleftharpoons \text{MIIa} \rightleftharpoons \text{MIIb} \rightleftharpoons \text{MIIbH}^+$$

Scheme 3: Equilibrium scheme of metarhodopsin states. Adapted from [56].

By adding Gα$_t$CT peptides to the spectroscopic titration assay, stabilizing interactions between peptides and specific metarhodopsin states can be investigated [56]. Following Scheme 4, the pH-dependent distribution of the metarhodopsin states was described. Addition of 10 mM peptide shifted the apparent pK_a to 9.1. A selective stabilization of the protonated MIIbH$^+$•Gα$_t$CT complex was found, extending over a much broader pH range of 5 to 9 than without Gα$_t$CT [56].

To investigate the effect of the Gα$_t$CT peptide binding on the receptor structure, a peptide binding spectrum (PBS) was devised [56]: Two difference spectra measured in the presence and absence of the Gα$_t$CT peptide are obtained; by subtracting one from the other, a double difference is calculated that defines the PBS. The subtraction cancels out the vibrational changes of the unbound receptor molecules, leaving the conformational changes to receptor and peptide upon complex formation [171, 172, 173]. Strong bands in the structurally sensitive amide II region of the Gα$_t$CT PBS indicate changes to the secondary structure of receptor and peptide [56].

$$\begin{array}{ccccccc} \text{MI} & \rightleftharpoons & \text{MIIa} & \rightleftharpoons & \text{MIIb} & \rightleftharpoons & \text{MIIbH}^+ \\ & & & & \downuparrows & & \downuparrows \\ & & & & \text{MIIb•G}\alpha_t\text{CT} & \rightleftharpoons & \text{MIIbH}^+\text{•G}\alpha_t\text{CT} \end{array}$$

Scheme 4: Equilibrium scheme of metarhodopsin states extended by Gα$_t$CT. Adapted from [56].

The crystal structures of RhR* and RhR*•Gα$_t$CT show a hydrogen bond network in ICL3 that links TM5 and TM6. It is comprised of the residues K231^{ICL3}, E247$^{6.30}$ and T251$^{6.34}$ and has been suggested as a determinant of the active conformation [1]. K231^{ICL3} is also part of the conserved Y$^{5.58}$(x)$_7$K(R)^{ICL3} motif. The role of K231^{ICL3} was previously investigated in a metadynamics simulation study [113]. There, the K231A mutant was used to prevent participation of this residue in hydrogen bonds and thus (potentially) weakening the link between TM5 and TM6. Moreover, energetic analysis of the metadynamics simulations suggests that the K231A mutation considerably affects the distribution of activated receptor conformations [113].

The FTIR spectroscopic study therefore obtained a difference spectrum of the K231A rhodopsin mutant and of wild-type (WT) rhodopsin at MIIbH$^+$-favoring conditions [56]. A double difference spectrum K231A minus WT was calculated to highlight the effect of

the K231A mutation. Remarkably, that double difference closely resembles the inverted Gα_tCT PBS. As the spectra describe conformational changes, their similarity suggests that the receptor conformation that is stabilized by Gα_tCT is in turn destabilized by the K231A mutation. This implies that the conformation stabilized by Gα_tCT is also present as a significant fraction in the ensemble of receptor conformations in the absence of Gα_tCT [56].

4.1.3 *Hierarchy of conformational states in RhR**

By employing FTIR spectroscopy and MD simulations of membrane embedded rhodopsin, the static view provided by X-ray crystal structures could be revised with the help of complementary dynamic information. Both, the FTIR and the MD analysis, suggest that without Gα_tCT the active receptor exists in an ensemble of states. The MD analysis indicates a flexible ICL3 that adopts many different conformations. The spectroscopic study established the receptor's conformational diversity and analysis of the K231A mutant could link that diversity to ICL3.

To discuss the role of the intrinsically unstructured regions, we will integrate the results on ICL3 and Gα_tCT dynamics into the concept of protein dynamics developed by Frauenfelder and co-workers [174, 175]. The concept establishes a hierarchy of three conformational states by distinguishing between the exchange rates. The top tier (*tier 0*) consists of protein states that exchange on the microsecond timescale. For RhR* the equilibrium of metarhodopsin states MI, MIIa, MIIb and MIIbH$^+$ builds up *tier 0* (Scheme 3). The protein states in turn are each an ensemble of taxonomic sub-states of *tier 1* exchanging with nanosecond lifetimes. Further sub-division of the taxonomic sub-states yields the statistical sub-states of *tier 2* describing fluctuations that change on the picosecond timescale [176]. These relatively fast conversions imply energy barriers of less than 1 kT, which results in fluctuating behavior at physiological temperatures that can be readily observed in MD simulations. Cluster analysis of RhR and Ops* simulations shows ICL3 conformations exchanging with lifetimes of 10 ns to 100 ns, indicating fluctuations within taxonomic sub-states (Figure 11A, B). Within a single taxonomic sub-state, the backbone RMSF describes its statistical sub-states, e. g. for ICL3 in simulations of RhR* stabilized by Gα_tCT (Figure 11C).

A salient insight from the hierarchy of conformational states is that transitions within a lower *tier* can facilitate transitions within a higher *tier* and, especially, that taxonomic transitions allow for fast protein state transitions [52], which will be important for a functional interpretation of the FTIR and MD results.

4.1.4 *Gα_tCT binding stabilizes a single sub-state of RhR**

The extreme C-terminus of the Gα subunit becomes structured upon binding RhR*. In X-ray crystallography structures of nucleotide bound G$\alpha_{i/t}\beta\gamma$ [64], G$\alpha_i\beta\gamma$ [177] (close homologue of G_t) or Gα_s [65], the GαCT is not resolved. Solution NMR structures show that the Gα_tCT peptide forms a capped α-helix when bound to active rhodopsin [68, 69]. Upon binding of the peptide, the structure of RhR* also exhibits changes as shown by FTIR spectroscopy [173]. However, crystal structures of RhR* do not show major differences between structures with or without peptide [31, 30, 66]. This discrepancy likely stems from the constraints imposed by the crystallization conditions, in particular by the crystal lattice. All structures of RhR* feature extensive crystal contacts of ICL3 regardless of binding partner or lack thereof [30, 31, 67, 66, 167, 178, 179, 75]. Note that the intrinsic flexibility of ICL3 is a major obstacle to successful structure determination of GPCRs in general and many structures were resolved with a fused T4-lysozyme in place of the ICL3 [180].

Our MD simulations support the notion that RhR* features an ensemble of ICL3 conformations, exhibiting much more flexibility than the other intracellular loops and helix 8 (Figure 10). As the simulations were based on protonated R*H$^+$ (MIIbH$^+$ or Ops*H$^+$) structures, the observed fluctuations showing different bent and helical conformations of ICL3 can be assigned to the R*H$^+$ protein state. Furthermore, the conformations within the ICL3 ensemble exchange on timescales of 10^{-8} s (Figure 11B), which puts them in the realm of taxonomic sub-states (*tier 1*). Our analysis agrees with an EPR analysis on the α-helical content of ICL3 notably in the aspect that flexibility of ICL3 is shifted to the TM5 side [162]. Also, the FTIR study complementing the MD simulations finds features in the difference spectra after hydrogen/deuterium exchange, which suggest backbone structuring after Gα_tCT peptide binding [56].

The spectroscopic and *in silico* data strongly suggest that ICL3 constitutes an intrinsically unstructured region that populates the taxonomic sub-states in the R*H$^+$ protein state. Adding a Gα_tCT peptide drastically changes the structure of ICL3 by altering the underlying conformational energy landscape from a rugged but relatively flat shape into a rather deep well (Figure 22) that corresponds to a largely α-helical ICL3 conformation (Figure 10C). This stabilization of one specific taxonomic sub-state of R*H$^+$ is apparent from distinct features in the peptide-binding spectrum [56]. The involvement of ICL3 in the peptide binding can be deduced from the destabilizing effect of the K231A mutation as shown by the FTIR analysis [56]. ICL3 structuring or decrease in solvent accessibility upon complex formation was also suggested by a hydrogen-deuterium exchange study [84]. Notably, it had been proposed based on sequence analysis of GPCRs

that ICL3 forms an intrinsically unstructured region that gets structured through complex formation of receptor and G protein [61].

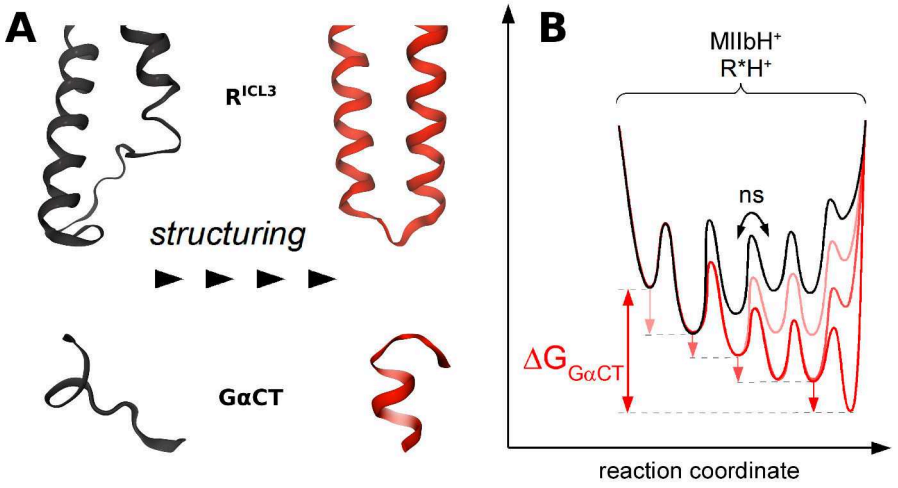

Figure 22: Mutual structuring of the receptor G protein coupling interface. $G\alpha_tCT$ specifically binds to the protonated active rhodopsin state ($MIIbH^+$, Ops^*H^+), thereby stabilizing a single taxonomic sub-state. (A) Upon $RhR^*\bullet G\alpha_tCT$ complex formation, both $G\alpha_tCT$ (bottom) and ICL3 (top) become structured (red cartoons) through mutual adjustment of the initially unstructured binding partners (black cartoon). Shown are the the unstructured ICL3 from RhR (1U19), the structured ICL3 from RhR* (3PXO), an unstructured $G\alpha_tCT$ (obtained by a MD simulation of $G\alpha_tCT$ in water) and the structured $G\alpha_tCT$ found in co-crystals of RhR* and $G\alpha_tCT$ (3PQR). (B) Schematic display of the mutual structuring of $G\alpha_tCT$ and ICL3 along a reaction coordinate describing ICL3 flexibility from unstructured (left) to structured (right). In the absence of $G\alpha_tCT$, ICL3 is more likely to be unstructured than structured (back energy landscape). Upon addition of $G\alpha_tCT$, the energy landscape is stepwise adjusted by interactions between $G\alpha_tCT$ and ICL3 until the final complex is formed. Structuring of $G\alpha_tCT$ at the coupling interface gradually makes a structured ICL3 conformation more favorable (red arrows and red energy landscape). Note that the individual energy levels and energy barriers are only schematic and do not represent experimentally obtained values. *The figure and legend are adapted from [56]*

4.1.5 *Binding of $G\alpha_tCT$ to the active receptor by mutual adjustment*

In both binding partners, receptor and G protein, a main part of their coupling interface is disordered until complex formation. The receptor exhibits a flexible ICL3 and the GαCT is unstructured [68, 64]. Given their disordered property, complex formation purely by conformational selection would be statistically improbable and thus extremely slow: Consider the multitude of taxonomic sub-states in a disordered region and how likely it would be to select the correct sub-state by chance. However, since active rhodopsin can activate up

to 10^3 G proteins per second and more under optimal conditions [44], we propose (based on the data produced in this study) a mutual adjustment coupling mechanism that starts after receptor and GαCT are in proximity.

Formation of an encounter complex [181] between activated receptor and GαCT is likely facilitated by the long-range charge potential (~1/r) of $R135^{3.50}$ and other charged amino acids exposed at the receptor binding crevice (Figure 17F). The encounter complex changes the local environment of both Gα_tCT and ICL3 by bringing them together such that short-range hydrophobic interactions (~1/r6) may build up. This would promote helix formation in GαCT and ICL3 through the associated expulsion of water molecules.

To conclude, we propose the following mutual structuring scheme as a mechanism allowing a fast and precise signal transduction from receptor to G protein. Once in contact, the ensembles of GαCT and ICL3 conformations become coupled. Then, a mechanism of stepwise complex formation takes over, comprising several stages of mutual adjustment [103] that shift the population of taxonomic sub-states towards the final complex. This shift corresponds to moving along a reaction coordinate in an energy landscape that is gradually altered as a result of the ongoing interactions between GαCT and ICL3 (Figure 22). With each step towards the final complex, more specific conformations along the reaction coordinate are selected. This stepwise progression works by transient population of distinct taxonomic sub-states exchanging within nanoseconds. The resulting binding funnel [182] has a large opening, accommodating a multitude of encounter complexes to start the formation of the final complex. By initially requiring only a structurally rather unspecific encounter-complex, formation of the final complex becomes more probable than in a purely conformational selection scheme.

4.2 TM6 POSITION AND β_2AR* G PROTEIN COUPLING

The β_2-adrenoceptor can couple to the G protein families G_s and G_i [4]. Moreover, the receptor must be able to distinguish between G_s and G_i, as there are known ligands, which can bias the receptor to prefer coupling to one G protein family [163]. The structural mechanism by which β_2AR* specifically selects G_s or G_i is unknown. However, spectroscopic studies have linked the dynamics of TM6 to G protein versus arrestin selectivity [27, 26].

In this study, we explored the dynamics of TM6 in active β_2AR*, employing MD simulations to address the issue of G_s versus G_i selectivity. Motivated by the differing TM6 conformations found in β_2AR*•G_s and RhR*•$G\alpha_t$CT, we investigated the TM6 behavior after removal of the constraining binding partner (i. e. G_s or $G\alpha_t$CT, respectively). We find that the TM6 in β_2AR* populates a larger conformational space than suggested by the β_2AR*•G_s crystal structure. One of the observed TM6 conformations looks very similar to the TM6 conformation of the crystal structures of RhR*•$G\alpha_t$CT. Extensive simulations and energy calculations then indeed suggest that $G\alpha_i$CT can stabilize β_2AR*•$G\alpha_i$CT complex that again looks very similar to the RhR*•$G\alpha_t$CT crystal structure.

Additional simulations indicate that a RhR*•$G\alpha_s$CT complex is not feasible. This is in line with a structural comparison of the β_2AR*•G_s and RhR*•$G\alpha_t$CT crystal structures, indicating that specific interaction of the bulkier $G\alpha_s$CT with $R^{3.50}$ requires a larger TM6 outward tilt than the slimmer $G\alpha_t$CT. In Section 4.1, we already studied the flexibility of ICL3, the loop connecting TM5 with TM6, in the rhodopsin system. The different roles of TM6 and ICL3 in receptor activation and G protein coupling will be discussed.

4.2.1 *Conformational heterogeneity of TM6 in active β_2AR**

In MD simulations, we observe a strong structural heterogeneity for the active GPCR β_2AR* with bound agonist and in the absence of intracellular interaction partners. We find different TM6 tilts broadly distributed, creating a conformational space of the active β_2AR*. The TM6 tilt distributions exhibit two maxima, one at 23 Å and the other at 28 Å. The G_s coupled state [32] is well represented by the pronounced maximum at 28 Å. The maximum at 23 Å, however, does not correspond to any structurally resolved β_2AR conformation. But it indicates a second β_2AR* conformation with a TM6 tilt, very similar to the 23.2 Å observed in the X-ray structure of the RhR*•$G\alpha_t$CT complex [31].

Our observations of structural heterogeneity are in accordance with recent studies that provide experimental evidence of structural heterogeneity in active β_2AR*. Two major conformational states of TM6

and TM7 in β_2AR* were observed in an NMR study with site-specific cysteine labels and different ligands [27]. Similarly, different conformational states of β_2AR* depending on ligand and intracellular binding partner were found in a study combining NMR spectroscopy with methionine labels and µs timescale MD simulations [55]. Another NMR study with a cysteine label on the intracellular side of TM6 observed multiple β_2AR conformations depending on the type of ligand and the presence of a G protein mimetic nanobody [164].

Moreover, we find the α-helical content of TM6 decreasing after removal of G_s and increasing upon complex formation (Figure A6A) with $G\alpha_i$CT. In simulations of β_2AR*•G_s, the α-helical content of TM6 remains unchanged. These observation for the β_2AR* system are very similar to our finding for the RhR* system, where we focused on the ICL3, linking TM5 and TM6. The similarity of α-helical content of TM6 in β_2AR* and RhR* is also remarkable insofar as TM6 in β_2-adrenoceptor is much less restricted by its ICL3 than in rhodopsin with a much shorter ICL3. In our combined FTIR and MD study, we find that in RhR* the ICL3 exists in an disordered ensemble and upon G protein binding adopts a single ordered conformation (see Section 4.1 and [56]). Both systems seem to go from disordered, more flexible conformations to a single specific conformation in the final complex. In Section 4.1, we argue that such an ordering process allows for fast and precise binding events.

4.2.2 *Stabilization of a more closed β_2AR* conformation by $G\alpha_i$CT*

In MD simulations of β_2AR* with $G\alpha_i$CT, a close homologue of $G\alpha_t$CT, we observe a defined β_2AR*•$G\alpha_i$CT complex that is very similar to the RhR*•$G\alpha_t$CT complex determined by X-ray crystallography (Figure A7C). This β_2AR*•$G\alpha_i$CT complex exhibits a more inward tilted TM6 than the β_2AR*•G_s complex (Figure 15). As we find this more inward tilted TM6 already present in the conformational space of the active β_2AR*, it is likely that the $G\alpha_i$CT peptide selects it from the equilibrium and stabilizes it. In contrast, we did not observe a β_2AR*•G_s or β_2AR*•$G\alpha_s$CT complex with an inward tilted TM6 in our simulations. Additionally, simulations of a putative RhR*•$G\alpha_s$CT peptide complex proved unstable (Figure A2F), which is in line with G_s not coupling with RhR* [165] and the absence of a G_s signaling pathway in the RhR system [12].

Since the difference in TM6 outward tilt is the major structural difference between β_2AR*•G_s and β_2AR*•$G\alpha_i$CT or RhR*•$G\alpha_t$CT, we assume that TM6 plays a significant role in determining whether a receptor or receptor conformation can or can not bind to G_s or $G\alpha_{i/t}$. Our structural comparison of the binding crevice in β_2AR*•$G\alpha_s\beta\gamma$ and in RhR*•$G\alpha_t$CT (Section 3.2) suggests that specific interaction of the bulkier $G\alpha_s$CT with $R^{3.50}$ needs a larger outward-tilted TM6,

resulting in a more open binding crevice. The MD simulations corroborate this interpretation of the complex structures. Taken together, our observations suggest that G_s and $G\alpha_{i/t}$ stabilize different receptor conformations. An analysis of the interactions between GαCT and receptor finds the two binding modes governed by different sets of specific interactions by which the GαCT imprint their own shape onto the receptor binding crevice (Figure 17).

Moreover, our analysis of G_i vs G_s interactions with the active receptors β_2AR* and RhR* provides a structural explanation of previous mutational studies. A chimera of $G\alpha_s$ and $G\alpha_t$ with only 11 C-terminal residues from $G\alpha_t$ could successfully bind to RhR*, resulting in potent activation [73]. Alanine mutants at the GαCT identified C347 and G348 as essential for activation by RhR*. The corresponding positions in G_i and G_s are Y391, E392 and C351, G352. The importance of the positions is reflected in the differential interactions of G_s and G_i we observe in the simulations. These are Y391 to $R^{3.50}$, E392 to TM6/H8 for G_s (Figure 17A) or N347 with ICL2, C351 with $R^{3.50}$, and D350 with ICL1 for G_i (Figure 17B).

β_2AR•Gα_iCT complex formation*

To characterize the observed β_2AR*•Gα_iCT complex formation in terms of energy differences (ΔG), we performed umbrella sampling calculations. We find while β_2AR*•Gα_iCT complex formation is retarded by an energy barrier, once this is overcome, TM6 inward movement commences in a continuous downhill reaction. Crossing over the energy barrier seems to correlate with the increase of the β_2AR* Gα_iCT19 interaction surface accompanying the TM6 inward movement (Figure 16A,B). This suggests that the corresponding water displacement plays an important role in β_2AR*•Gα_iCT complex formation, indicating an entropic contribution to the binding energy [183].

4.2.3 *Extrapolation of β_2AR*•Gα_iCT to a β_2AR*•G$\alpha_i\beta\gamma$ model*

The β_2AR*•G_s complex [32] shows the arrangement of β_2AR* with an activated G_s. Defining features of G_s in this complex are a rotated and translated α-helix in $G\alpha_s$ together with a displacement of the $G\alpha_s$ all-helical domain. By employing the G_s from β_2AR*•G_s complex to extrapolate from the Gα_iCT 19 peptide to the G_i holoprotein, the defining features of an active G could be retained in a putative β_2AR*•G$\alpha_i\beta\gamma$ complex. Moreover, the such modeled β_2AR*•G$\alpha_i\beta\gamma$ complex yields a feasible arrangement of β_2AR* and G_i without internal clashes (Figure A7A) or unexpected membrane contacts/insertions.

In the context of G protein activation, GDP release is mainly triggered (apart from intrinsic activity, if any) by complex formation with an active GPCR. In Section 4.3, we propose a mechanistic role

of the α5 helix displacement for GDP release in the β_2AR*•G_s and RhR*•G_t systems. Also, the model of β_2RhR*•$G\alpha_i\beta\gamma$ by Alexander, Preininger, Kaya, Stein, Hamm, and Meiler [87] features the same general $G\alpha_i$CT binding mode as in our β_2AR*•$G\alpha_i$CT complex. Given the high degree of similarity between the β_2AR*•$G\alpha_i$CT and the RhR*•$G\alpha_t$CT complexes, the proposed helix switch should be applicable to the β_2AR*•G_i system as well.

4.2.4 *TM6 tilts, ICL3 flexibility and G protein binding modes*

The GPCR structures resolved so far show the same overall 7-TM fold, even in distantly related receptors from different families [34]. A notable difference with regard to the TM bundle is the position of TM6. Generally the TM6 positions fall into two clusters, one that is associated with the inactive receptor state and another representing the active receptor, capable of binding a G protein. In the inactive receptor, the intracellular part of TM6 is buried within the TM bundle and directly contacting TM3 [36, 37, 38]. Upon receptor activation, TM6 tilts outward, increasing the distance between TM3 and the intracellular part of TM6 by more than 6 Å. This hallmark event of GPCR activation creates a binding crevice at the intracellular side to which a G protein can bind [1].

While there are now structures of inactive receptors for most GPCR families, structural data on active receptors is still sparse [2, 3]. However, beginning with TM3-TM6 distance measurements by a SDSL EPR study comparing inactive to active rhodopsin, experimental evidence was provided for a substantial TM6 movement upon activation [28, 29]. Later, structures of RhR*, β_2AR* and M2* resolved by X-ray crystallography showed the active receptors in full atomistic detail [30, 32, 33]. These structures also exhibit a range of TM6 positions, that depend on the system, the interaction partner or the experimental conditions [33, 3]. Still, all these TM6 positions create an intracellular binding crevice, thus suggesting an ensemble of TM6 conformations in active GPCRs.

Investigating the role and the source of the ensemble of TM6 conformations in active GPCRs and their complexes with intracellular binding partners is an ongoing research effort. Initially, it had was proposed that the β_2AR*•G_s structure [60] shows the productive signaling complex, whereas the RhR*•$G\alpha_t$CT structure [31] depicts part of the signaling complex in an earlier state, closer to the first encounter of receptor and G protein [184, 60]. However, this interpretation ignores that the two structures are complexes derived from different receptors and different G proteins. The β_2AR*•G_s structure shows a cation-π interaction between Y391 and $R^{3.50}$, whereas in RhR*•$G\alpha_t$CT a backbone hydrogen bond between C347 and $R^{3.50}$ connects the Gα reverse turn with receptor binding crevice (Figure 14A,B). The

cation-π interaction is not possible for the $G\alpha_{i/t}$ family as it lacks a tyrosine in the $G\alpha$ reverse turn. Numerous spectroscopic studies have now shown the influence of binding partners and ligands on the β_2AR* conformations, especially on the TM6 conformation [27, 55, 164]. In Section 4.2.1, we propose that β_2AR* employs different TM6 positions to accommodate the two different G proteins G_s and G_i β_2AR* is capable of binding and activating. A crystal structure of a β_2AR*•G_i or RhR*•G_t holocomplex would go a long way tracking down differences in G_s versus $G\alpha_{i/t}$ binding modes and explaining how GPCRs accommodate different G proteins.

ICL3 flexibility

Closely related to the issue of TM6 conformations is the structure and flexibility of the ICL3 linking TM5 and TM6. In Section 4.1 we discussed the ICL3 behavior in the RhR G_t system. There we provide evidence that the ICL3 is a disordered region that becomes structured upon binding to $G\alpha_t$CT. The intracellular ends of TM5 and TM6 exhibit order to disorder transitions, blurring the exact start and end of the ICL3. We further propose that the binding is driven by a mutual structuring process involving ICL3 and $G\alpha_t$CT, allowing for fast and precise signaling. In the β_2AR system investigated in this section, the ICL3 is much longer than in RhR and not resolved in the available crystal structures [40, 60, 32]. The lack of structural data on β_2AR ICL3 suggests it is disordered, which corroborates similar observations in our simulations of β_2AR* and β_2AR*•G_s. The intracellular domains of GPCRs, especially ICL3, are generally predicted to contain intrinsically disordered regions for all GPCR classes, not just for members of class A as investigated here, and are implicated in fine-tuning receptor activity and mediating interactions with their effector proteins, for example G proteins [61, 62]. When we focused our analysis on the intracellular end of TM6 instead of the full ICL3, we made similar observations as in our RhR ICL3 study. Without the G_s protein, some disordering can be observed at the intracellular TM6 end of β_2AR* (Figure A1). Notably, in one simulation of β_2AR*•$G\alpha_i$CT where TM6 moves inward, the full ICL3 adopts an ordered conformation that it maintains for the remainder of the simulation (Figure A6).

4.2.5 *Structural mechanism of G protein selection by β_2AR**

We investigated how G_i might bind to β_2AR* and what the differences are compared to G_s binding. For this purpose, we performed MD simulations of β_2AR* with $G\alpha_i$CT and $G_s/G\alpha_s$CT. Our results add to the recent findings suggesting that GPCR signaling is based on an ensemble of active receptor conformations [27, 55, 164] which are influenced by a variety of extracellular ligands and intracellular

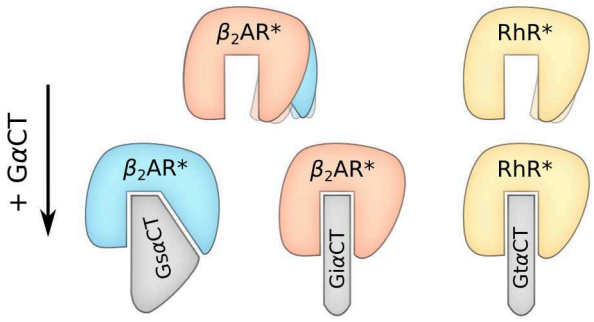

Figure 23: β_2AR and RhR specificity scheme. Upon binding of GαCT (+GαCT), a single receptor conformation is adopted. The equilibrium of active receptor states includes in the case of β_2AR* conformations selective for $G\alpha_s$CT and $G\alpha_i$CT. In RhR*, the conformational space is more specialized and tuned to support binding only to $G\alpha_t$CT. Notably, the binding modes of β_2AR*•$G\alpha_i$CT and RhR*•$G\alpha_t$CT are very similar, presumably due the very high sequence similarity of $G\alpha_i$CT and $G\alpha_t$CT.

proteins [26, 185, 186]. Ligands specifically shift the equilibrium of receptor conformations — and not just towards the inactive or active state (i. e. as inverse agonists or agonists do respectively) but also towards active receptor conformations favoring specific effector proteins [163].

In this study, we employed the Gα C-termini of G_s or G_i to investigate the interactions with the active β_2AR*. GαCT is known to be a main interaction site with the GPCR governing coupling fidelity and specificity [70, 71, 72, 73, 74]. Thus, $G\alpha_s$CT and $G\alpha_i$CT represent key sites for interactions with the active β_2AR*, and our MD simulations indeed show them stabilizing distinct active receptor conformations. In the simulations, $G\alpha_s$CT maintains the β_2AR* conformation from the X-ray structure β_2AR*•G_s complex, which is also maintained in simulations with G_s. Moreover, simulations of β_2AR* show an ensemble of conformations from which $G\alpha_i$CT can stabilize a β_2AR* receptor conformation resembling RhR* (Figure 15). More precisely, the active receptor exhibits intrinsic TM6 flexibility spanning a conformational space from which the G protein selects a specific conformation for productive interaction and signal transfer. The Gα C-terminus is often sufficient to control the binding specificity by presenting just a small number of crucial residues (Figure 17 and ref [73]).

Structural analyses of β_2AR*•G_s and RhR as well as simulations of RhR*•$G\alpha_s$CT together with the putative β_2AR*•$G\alpha_i$CT complex suggest β_2AR* can adopt a RhR*-like conformation that excludes G_s from binding and being activated. Thus, by inducing a RhR*-like conformation of β_2AR*, ligands may switch β_2AR* to a receptor conformation that only activates G_i but not G_s [163]. To summarize, the ensemble of specific R* conformations constitutes a control mechanism for adjust which signaling pathway prevails in the intracellular

network. In addition to the availability and affinity of different G proteins, shifts within the ensemble of active receptor conformations may favor or abolish productive interaction and alter signal efficiency.

4.3 ROLE OF R* IN G α5 DISPLACEMENT

Signal transfer from GPCRs to G proteins has to bridge a distance of 30 Å to get from the receptor binding crevice to the GDP binding pocket. The α5 helix from the Gα subunit provides a direct connection between the GDP binding pocket and the coupling interface between receptor and G protein [32] which has been linked to GDP release [87]. Upon R*•G complex formation, the α5 is displaced, particularly rotated [80, 32]. Here we investigated the role of α5 and its displacement to elucidate the events leading to GDP release and thus to signal transduction.

For signal transfer, a series of intermediate states is occupied by receptors and G proteins reflecting an intricate interaction network that governs the complex formation required for GDP release. In these intermediates, participation of a number of protein domains is required in a sequential manner for the stepwise adoption of a productive complex [83, 48, 43, 56, 1, 187, 170, 188, 65, 82] (Figure 24). A number of crystal structures of the inactive and active forms of receptors [34] as well as G proteins and their complexes [76, 32] are available. Given the experimental condition in which they were resolved, these structures reflect stable states at fixed points in their series of intermediates and likely states at endpoints. A clear example of such a state is the β_2AR*•G_s complex, a GDP free state exhibiting the defining features of an active receptor G protein complex.

Considering the central position of the α5 helix during complex formation, we wanted to learn more about how and when the rotation of α5 of the α-subunit of G_s and G_t occurs within the binding crevice of β_2AR* and RhR*. For that, we created a model of the R*•G[GDP] intermediate, employing flexible docking analysis and validated it against available structural data (Figure 24). By starting MD simulations from this R*•G[GDP] intermediate, we obtained a dynamic view describing the effect the receptor has on the α5 helix. To keep the computations feasible, we studied the dynamics of $G\alpha_s$CT and $G\alpha_t$CT, derived from the far C- terminus of α5, in lieu of the complete G protein. With this setup, we tested whether the α5 helix can rotate given only the interactions of GαCT with R*.

4.3.1 *Characterization of the R*•G[GDP] intermediate*

First experimental evidence for a R*•G[GDP] intermediate was provided by a study on the kinetics of Gt activation in disc membranes [44, 78]. The R*•G[GDP] intermediate's defining features is that GDP is still bound in the nucleotide binding site of the Gα subunit after complex formation with the active receptor. Intermediates are generally seen as stable states linked by transition states, showing transitory or partial features of the final complex to which their reor-

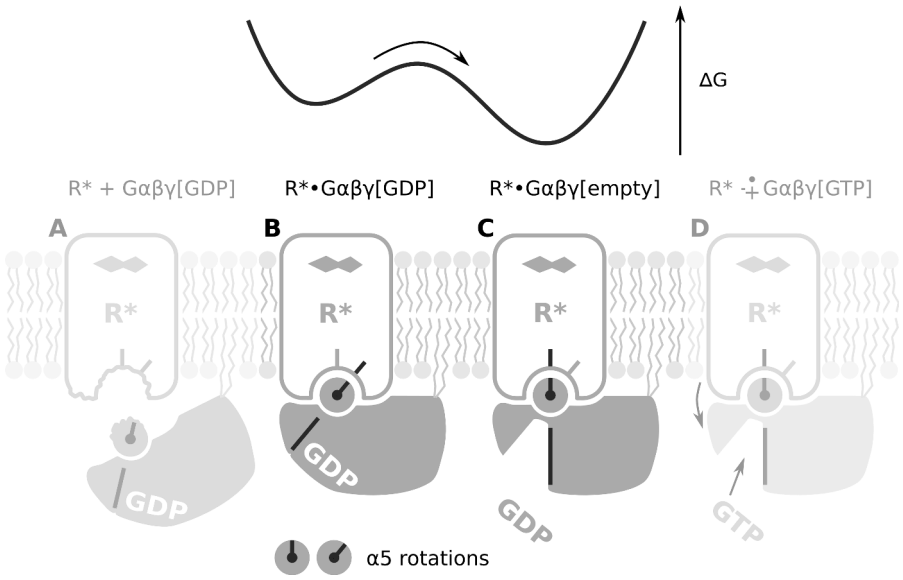

Figure 24: Role of the α5 helix in the interaction between R* and G that leads to nucleotide exchange. From left to right. (A) Membrane anchored Gαβγ[GDP] with an unstructured α5 C-terminus encounters R* with a partially unstructured cytoplasmic crevice. (B) Intermediate R*•Gαβγ[GDP] complex formed through mutual structuring of the α5 C-terminus and the R* cytoplasmic crevice. The α5 helix has not rotated compared to free R*• Gαβγ[GDP]. (C) Rotation of α5 has lowered the energy barrier separating R*•Gαβγ[GDP] from R*•Gαβγ[empty], resulting in GDP release. (D) Uptake of GTP completes the nucleotide exchange reaction accompanied by Gαβγ[GTP] uncoupling from R*. *The figure and legend are reproduced from [166]*

ganization eventually leads [189]. In the context of protein dynamics developed by Frauenfelder and co-workers [174, 175], intermediate states are considered protein states with microsecond, i. e. finite, lifetimes. Here we are interested in an intermediate complex formed by R* G[GDP] featuring a folded GαCT and ICL3 (or just the intracellular parts of TM5 and TM6 in case of a long ICL3). As argued in Section 4.1 and ref [56] for the signal transfer from RhR* to G_t, such an intermediate is formed as the result of a mutual structuring process by stepwise reduction of the conformational space.

By means of a flexible docking analysis, we obtained models of the R*•G[GDP] intermediate for the β_2AR* G_s and the RhR* G_t system. These intermediates show a double sandwich structure comprising αN/β2-β3 loop, ICL2, α5 and ICL3 (Figure 19A, for $G\alpha_t$CT see [78]), which is also seen in the β_2AR*•G_s complex, and exhibit the C-terminal reverse turn and a structured ICL3. In contrast to the β_2AR*•G_s complex, however, the α5 is not rotated/displaced in the intermediate.

Regarding the interaction of $G\alpha_t$CT with RhR* (see SI of [78]), we find a hydrogen bond in our intermediate from $G\alpha_t$CT to the conserved $P138^{3.57}$ cap interconnecting TM3 with ICL2. Mutations of

the ICL2 including the P138$^{3.57}$ have previously been used to create a RhR*•G_t complex that forms but is not productive [57]. Whether such an unproductive complex has GDP still bound is not clear, but interactions with the ICL2 seem to be crucial for successful signaling. For the β_2AR* G_s system, Sunahara and co-workers recently provided data that suggest GDP-bound $G\alpha_s$ couples to β_2AR* mainly through $G\alpha_s$CT [83]. In our intermediate $G\alpha_s$CT binds to β_2AR* by a hydrogen bond between E392 and R131$^{3.50}$ of TM3 employing the E(D)RY motif. Additional hydrogen bonds are found from $G\alpha_s$CT to ICL2 and 3 (Figure 19A, C).

Neither $G\alpha_s$CT nor $G\alpha_t$CT exhibit extended hydrophobic interactions in our intermediate states. As the full desolvation potential is not exploited, the absence of such interaction generally may be indicative of intermediate states. Here it also specifically reflects the weak coupling of GDP-bound $G\alpha_s$ to β_2AR* [83]. Accordingly, we observe in our MD simulations, a highly dynamic behavior of especially $G\alpha_s$CT but also of $G\alpha_t$CT in the intermediate position. As described in the next section, the GαCT peptides either unfold and diffuse away or switch very quickly to the position seen in the crystal structures of β_2AR*•G_s or RhR*•$G\alpha_t$CT, respectively. However, note that the very fast switching we observe is likely exaggerated due to our use of peptides in lieu of the complete G protein. We expect the intermediate to be stable for longer in the context of the full G protein with more restricting interactions when the GαCT is part of the Gα subunit and has no free N-terminus as with the $G\alpha_t$CT peptides.

4.3.2 *Switching of the α5 helix at the R* interface*

We performed 30 MD simulations each of the β_2AR* and of the RhR* systems. In both systems, about one third of the simulations feature a GαCT helix switch. Analysis of these simulations shows that the orientation and the key interactions present in the X-ray structures are eventually restored as a result of the GαCT helix switch (Figure 20). In the case of RhR* and $G\alpha_t$CT, our simulations can essentially reproduce the X-ray structure of RhR* with 11-mer $G\alpha_t$CT [31]. However, for β_2AR* and $G\alpha_s$CT, the available structural data is on the full β_2AR*•G_s complex, showing additional constraints on the $G\alpha_s$CT position as part of the whole Gα subunit. Still, we observe $G\alpha_s$CT restoring the orientation and key interactions seen in the crystal structure of the β_2AR*•G_s complex [32].

These observations show that $G\alpha_s$CT and $G\alpha_t$CT can transition from their intermediary position to the position and orientation representing the GDP-free complexes (Figure 20) without applying additional forces. Specifically, the interactions of the far C-terminus of α5 with the intracellular crevice of R* and the solvent are the only apparent forces we observe in the simulations. Thus, the ability of the α5

helix to switch at the R* interface seems to be an intrinsic feature of the coupling interface comprising GαCT and the intracellular crevice of R*. Following the theory of complex formation, transition from intermediate states to the final complex is accompanied by a reorganization of electrostatic interactions and desolvation of hydrophobic patches [189]. Accordingly, we observe changes in the hydrogen bond network (Figures 19A-D, A24, A25) and an expulsion of water leading to the formation of a hydrophobic contact patch at the binding interface of GαCT and ICL3 of R* (Figures A19, A28, A29).

In our proposed structural models, transitioning from the intermediate GDP-bound to the final GDP-free state, the rotation of the α5 helix inherently exposes two adjacent sites of the α5 helix to the intracellular binding crevice of R*. Thereby two different (but not distinct) interaction networks are involved during R*•G[GDP] formation and its transition to the GDP-free complex. Conceivably, these networks function to independently verify R* Gα C-terminus complementarity, that is they allow R*•G[empty] formation only after successful R*• G[GDP] formation. The initial interaction network for R*•G[GDP] would not require altering the position of α5. We propose (see next section) that rotation of the α5 helix triggers GDP release with formation of the second interaction network, yielding the R*•G[empty]. Thus the initial interaction network would be uncoupled from GDP release. Such a nuanced role of GαCT would be in line with and add to the already described complex role of GαCT in determining coupling specificity and efficiency [70, 74, 76].

Involvement of the highly conserved motifs at $R^{3.50}$ and at the $P^{3.57}$ cap of TM3 (Figures 19, A16, A17) in the hydrogen bond network between GαCT and R* indicates that the observed mechanism is applicable in similar form to other GPCR/G protein systems.

4.3.3 *Role of the α5 helix switch for GDP release*

The present study describes a structural mechanism by which the α5 helix switch could occur and lead to the release of GDP from the nucleotide binding pocket (Figure 24). An α5 helix displacement is known to happen upon R*•G[empty] complex formation and is thus coupled to GDP release. Interconnecting the R* cytoplasmic crevice with the G nucleotide binding pocket, the α5 helix is a prime suspect for bridging that distance and transducing a signal [87, 80, 32, 78].

Here we propose that the intermediate R*•G[GDP] complex provides a structural scaffold in which the α5 helix can rotate while the receptor and the G protein (especially the Gα subunit) stay fixed [78]. The structural scaffold is composed of the elements of the double sandwich structure from Gα and R*: αN-β1 or β2-β3 loop, ICL2, α5, ICL3 and the membrane anchors. Given such a fixed frame, the α5 helix can propagate the signal to the GDP binding pocket by rotating

within the intracellular crevice of R*. In MD simulations of GαCT, we observe that the α5 helix switches resulting from interactions with at the R* interface alone, indicating the feasibility of the scaffold mechanism.

Signal propagation by α5 rotation presumably works through destabilization of the GDP binding pocket. Specifically, changes of the α5-β5 loop have been identified as one of the key elements in the course of GDP release [87]. While GDP is bound, the α5-β5 loop of the G protein is highly constrained, forming a short structural element with strong polar interactions to GDP. Due to the counter clockwise rotation of α5, the α5-β5 loop is eventually elongated as seen in the the X-ray structure of the β_2AR*•G_s complex. Moreover data from HDX experiments strongly suggest that this loop becomes more flexible in the β_2AR*•G_s complex [83]. Therefore, the α5 helix switch may facilitate GDP release by releasing the constraints on the α5-β5 loop and thus playing an integral part in the mechanism of receptor G protein coupling.

5 CONCLUSION AND PERSPECTIVE

5.1 DYNAMIC COUPLING INTERFACE OF GPCRS

The results presented here show that the GPCR coupling interface features highly dynamic regions that adapt to accommodate different intracellular binding partners or the same binding partner in different and highly specific conformations. However, before the receptor forms a productive signaling complex with its binding partner, an agonist needs to activate the receptor [1, 2, 3]. In case of rhodopsin, the covalently linked *11-cis* retinal acts as an inverse agonist which is isomerized by light absorption to the agonist *all-trans* retinal [1]. Other GPCRs — like the β_2-adrenoceptor— show considerable constitutive activity [47], without a bound agonist. Additionally, GPCR activation depends on pH [50, 169, 23] and lipid composition [24], emphasizing the importance of experimental conditions. Eventually, the active receptor presents itself as an equilibrium of active conformations that share major rearrangements of the transmembrane helix bundle compared to the inactive receptor state [28, 29]. The exact distribution of the active conformations is governed by specific agonists which may shift the equilibrium of active conformations towards specific conformations [18, 27] and thus ultimately regulate which cellular signaling pathway is activated [13].

But first, to form a signaling complex, the receptor and its intracellular binding partners need to find each other and form loosely coupled encounter complexes [181]. Electrostatic interactions are known to facilitate the formation of such encounter complexes [190]. The exposure of $R^{3.50}$ (from its buried state) during receptor activation, along with other positively charged residues at the intracellular receptor interface, may thus help to facilitate encounter. For rhodopsin, a sequential fit mechanism has been suggested based on the observation that the C-terminus of the γ-subunit (GγCT) can interact with earlier rhodopsin intermediates than the main interaction site GαCT [187, 191]. In the sequential fit mechanism, GγCT may help to form earlier encounter complexes that allow for faster formation of the signaling complexes. Arrestin, on the other hand, needs a phosphorylated receptor C-terminus in addition to an active receptor conformation [192]. Interaction between the *p44* splice-variant of arrestin and the receptor's C-terminus is also observed with the inactive receptor, allowing for some form of pre-complexes [193]. A general effect on complex formation would come from superstructures proposed for rhodopsin where receptor molecules form large ordered

structures [194, 195]. For G proteins, this suggests that encountering one receptor in the superstructure may speed up subsequent encounters of receptors due to the spatial proximity [196, 197].

Mutual structuring of ICL3 and GαCT

We propose that, following the creation of an encounter complex, a mutual structuring process leads to the formation of a specific complex. Both the receptor and its intracellular binding partners include an intrinsically unstructured region at sites known to be crucial for signaling complex formation [162, 68, 84]. The ICL3 of the receptor is known to be flexible and has often been replaced to facilitate crystallogenesis [180]. In the case of rhodopsin the ICL3 is so short that crystal contacts stabilize it in the available crystal structures of active conformations [31]. However, in this combined study employing FTIR spectroscopy and MD simulations we were able to describe the intrinsically disordered nature of ICL3 on a structural level [56]. Moreover, we identified MIIbH$^+$ as the specific receptor conformation to which Gα_tCT binds and in which the ICL3 becomes structured (Section 4.1). While the ICL3 from the receptor becomes structured, the counterpart at the intracellular binding partner becomes structured too. GαCT forms a continuous α-helix capped by a reverse turn motif only when bound to an active receptor [68]. The "finger loop" in arrestin is more flexible, exhibiting different conformations [198, 199]. A peptide derived from the "finger loop" (ArrFL) becomes structured similarly to Gα_tCT and binds to the same intracellular receptor as GαCT [75].

In MD simulations, we were able to observe destructuring events of ICL3, GαCT and ArrFL. But the reverse process remains to be elucidated. To observe the destructuring, it was sufficient to remove the respective binding partner. For the reverse process, i. e. the structuring, a more elaborate approach is needed. A suitable encounter complex (or better yet, an ensemble of encounter complexes) would be required to start simulations from — with the aim to observe structuring. Given the large degree of freedom within such encounter complexes [200], it is likely necessary to bias or simplify the simulation to increase the chance of observing structuring with adequate computational resources. One avenue could be to have one of the binding partners already structured, for example the ICL3, and then focus on GαCT folding within the receptor binding crevice environment. Another approach would be to start the simulations with the binding partners in ways that limit their degree of freedom and thus allowing extensive sampling.

In reality, mutual structuring will eventually take place, regardless of whether it is intractable for MD simulations or not, and lead to the formation of a specific complex [103]. In the case of the interaction between receptor and G protein, the first specific complex will be the R*•G[GDP] intermediate [44, 78]. As described earlier, however,

whether a productive signaling complex can be efficiently formed depends on the specific distribution of conformations in the equilibrium of active receptor conformations.

TM6 conformations and biased signaling

The intracellular binding partners of GPCRs may show different binding strength or no binding at all with respect to specific active receptor conformations [18, 27]. Thus the stabilization of specific active receptor conformations, by different agonists, can enhance (or inhibit) the formation of signaling complexes and through that provide a structural mechanism explaining the biased signaling capabilities of ligands [25]. Several spectroscopic studies have found TM6 exhibiting multiple conformations in β_2AR*, linked to the presence of different ligands and/or intracellular binding partners [27, 55, 164]. Here we employed MD simulations, after a detailed structural comparison, to find a putative β_2AR*•Gα_iCT complex with a TM6 conformation distinct from β_2AR*•G_s but similar to RhR*•Gα_tCT. Given the close relationship between G_i and G_t, the β_2AR*•Gα_iCT complex suggests that the TM6 conformation in R* depends on which G protein type it is bound to: a small outward tilt of TM6 is sufficient for G$\alpha_{i/t}$, but a large TM6 outward tilt is required for G_s. This interdependence of G protein and TM6 conformation readily provides an explanation for the disparity of TM6 outward tilts in the crystal structures of the β_2AR*•G_s and RhR*•Gα_tCT complexes [31, 32].

Now, assuming that there are two active receptor conformations with respect to TM6, one for G_i and on for G_s, implications for biased signaling are apparent. The β_2-adrenoceptor can activate both the stimulatory G_s and the inhibitory G_i protein signaling pathway [4]. Moreover, a β_2AR ligand is known that can selectively activate the G_i while silencing the G_s pathway [163]. In light of our β_2AR*•Gα_iCT complex with its specific TM6 conformation prohibiting G_s binding, we suggest that G_i selective ligands shift the equilibrium of active receptor states towards this G_i specific TM6 conformation. Thus, ligands may utilize TM6 to allow the receptor to distinguish between different G proteins. Similar mechanisms are directly conceivable for other structural elements in the receptor that undergo major structural rearrangements during activation like TM5 or ICL2 [53].

The structural link between ligands and changes at the intracellular binding crevice is only poorly understood. TM6 plays a major role in both receptor activation and in the equilibrium of active receptor states [29, 55]. Thus, it is conceivable that the connector ($I121^{3.40}$ and $F282^{6.44}$ in β_2AR) proposed by Dror, Arlow, Maragakis, Mildorf, Pan, Xu, Borhani, and Shaw [54] for receptor activation is also important for modifying the conformational equilibrium of active receptor states through biased ligands. Initial MD simulation studies show a ligand dependence for the receptor activation pathway [116] and the

subsequent conformational equilibrium of active receptor states [201]. Given that TM6 and its large movements are so crucial, it would be valuable to make TM6 the center of a MD simulations study: by sampling the conformational space of TM6 in the equilibrium of active receptors, the boundaries of TM6 movement could be described and involved residues inferred. Moreover, performing such a study with different ligands would yield insights into the mechanism by which ligands act on the conformational space of TM6.

α5-helix rotation during GDP release

Once the receptor is activated, a complex with an intracellular binding partner can be formed, depending on the equilibrium of active receptor conformations. Eventually, a productive signaling complex will be formed that relays the signal to downstream effectors [6, 13]. However, for the interaction between receptor and G protein, the first specific complex will be — presumably via a mutual structuring process — the R*•G[GDP] intermediate. In this study, we focused on the next step of GDP release, to which the displacement of the α5-helix is central [80, 78, 83]. The α5 helix bridges the distance of 30 Å between receptor nucleotide binding pocket, allowing allosteric signal transmission. Upon formation of the R*•G[GDP] intermediate and the following transition to the nucleotide-free R*•Gαβγ[empty] complex, the α5-helix from the Gα-subunit is displaced. Then, the nucleotide exchange (GTP for GDP) is completed by proceeding through R*•Gαβγ[empty] and the following separation of Gαβγ[GTP] from R* [76].

Given its place as the first specific complex during nucleotide exchange, we suggest that the R*•G[GDP] intermediate provides a structural scaffold in which the displacement of the α5-helix can occur. Since the α5-helix provides an — at least allosteric, if not direct — connection between receptor and GDP binding pocket, its rotation is likely involved in GDP release [80, 88, 89, 90, 78]. Here we propose that the rotation of the α5-helix proceeds after formation of the R*•G[GDP] intermediate, in a fixed frame provided by additional interaction between receptor and G protein. In MD simulations of $G\alpha_s$CT and $G\alpha_t$CT peptides, we observed intrinsic switching of the α5 helix at the interface of β_2AR* and RhR*, respectively, indicating the feasibility of the scaffold-based mechanism. Starting from our R*•G[GDP] intermediate and employing sufficient computing resources, we could extend the GαCT peptide-approach to the full G protein. This will allow us to directly investigate the effect the rotation of the α5-helix within the R*•G[GDP]-provided scaffold has on the GDP binding pocket.

A

APPENDIX

System	Simulations	Length
RhR	1× 400 ns, 2× 200 ns	0.8 µs
RhR*	1× 400 ns, 4× 200 ns	1.2 µs
RhR*•Gα_tCT19	5× 200 ns	1.0 µs
RhR*•Gα_tCT11	1× 400 ns, 4× 200 ns	1.2 µs
RhR*•Gα_tCT11 intermediate	30× 100 ns	3.0 µs
RhR*•Gα_sCT19	5× 200 ns	1.0 µs
RhR*•Gα_sCT19 intermediate	1× 400 ns, 3× 200 ns	1.0 µs
Ops*	1× 400 ns, 4× 200 ns	1.2 µs
Ops*•Gα_tCT11	3× 400 ns, 7× 200 ns	2.6 µs
β_2AR*	1× 400 ns, 4× 200 ns	1.2 µs
β_2AR*•G$\alpha_s\beta\gamma$	5× 200 ns	1.0 µs
β_2AR*•Gα_sCT19	1× 400 ns, 4× 200 ns	1.2 µs
β_2AR*•Gα_sCT11	10× 500 ns	5.0 µs
β_2AR*•Gα_sCT11 intermediate	30× 200 ns	6.0 µs
β_2AR*•Gα_iCT19	3× 600 ns, 8× 200 ns, 10× 100 ns	4.4 µs
β_2AR*•Gα_sCT19 intermediate	1× 400 ns, 5× 200 ns	1.4 µs
β_2AR*•Gα_iCT19 umbrella sampling	36× 200 ns	7.2 µs
		40.4 µs

Table A1: Overview of performed MD simulations. Each row lists the system, the length and count of simulations, and the overall per system simulation time. The total simulation time is 40.4 µs.

β_2AR	Conservation	Location
T66	T 46%, S 16%, N 7%, K 4%, R 4%, E 3%, Q 3%, A 3%	ICL1
N69	Y 59%, N 15%, D 4%, T 3%, S 2%, F 2%, A 2%	ICL1/TM2
R131$^{3.50}$	R 89%, C 3%, H 2%	TM3
P138$^{3.57}$	P 86%, A 6%	TM3
D331	E 30%, D 24%, Q 10%, K 6%, N 5%, T 3%, S 3%, R 3%, H 2%	H8

Table A2: Conservation of residues from the binding cavity of R* important for stabilization of the R*•G[GDP] intermediate. Conservation was calculated from the alignment of all sequences from the UniProt data base (`http://www.uniprot.org/`, release 2012_07) which belong to the *g protein coupled receptor 1 family* and are in the human genome. Sequence alignment was performed with the Clustal Omega program. Only amino acids observed in at least 2% of the aligned sequences are listed. *The figure and legend are reproduced from [166]*

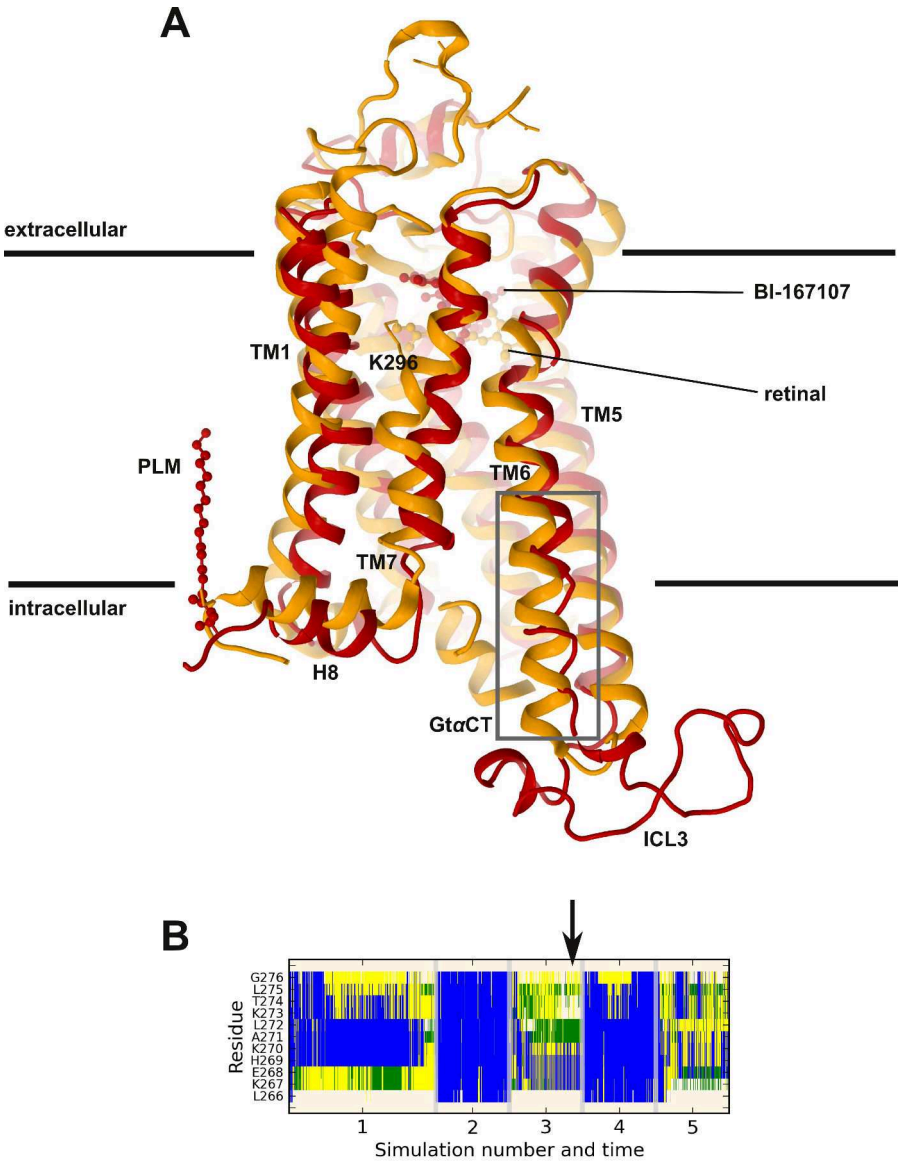

Figure A1: (A) Superposition of the RhR*•Gα_tCT crystal structure (orange, PDB entry 3PQR) with a snapshot from the MD simulation of uncomplexed β_2AR* (red). Note the similarity of the TM6 tilt at the intracellular side. The gray box denotes the part of TM6 analyzed in B. (B) Time series showing the secondary structure assignment of the intracellular TM6 region of β_2AR* as calculated with DSSP [155]. The observed structure types are α-helix (blue), 3_{10} helix (gray), turn (yellow), coil (white) or bend (green). The black arrow denotes from which simulation and at what time the β_2AR* structure shown in (A) was taken. The plot shows the data of multiple MD simulations, each between 200 ns and 400 ns long. *The figure and legend are reproduced from [101]*

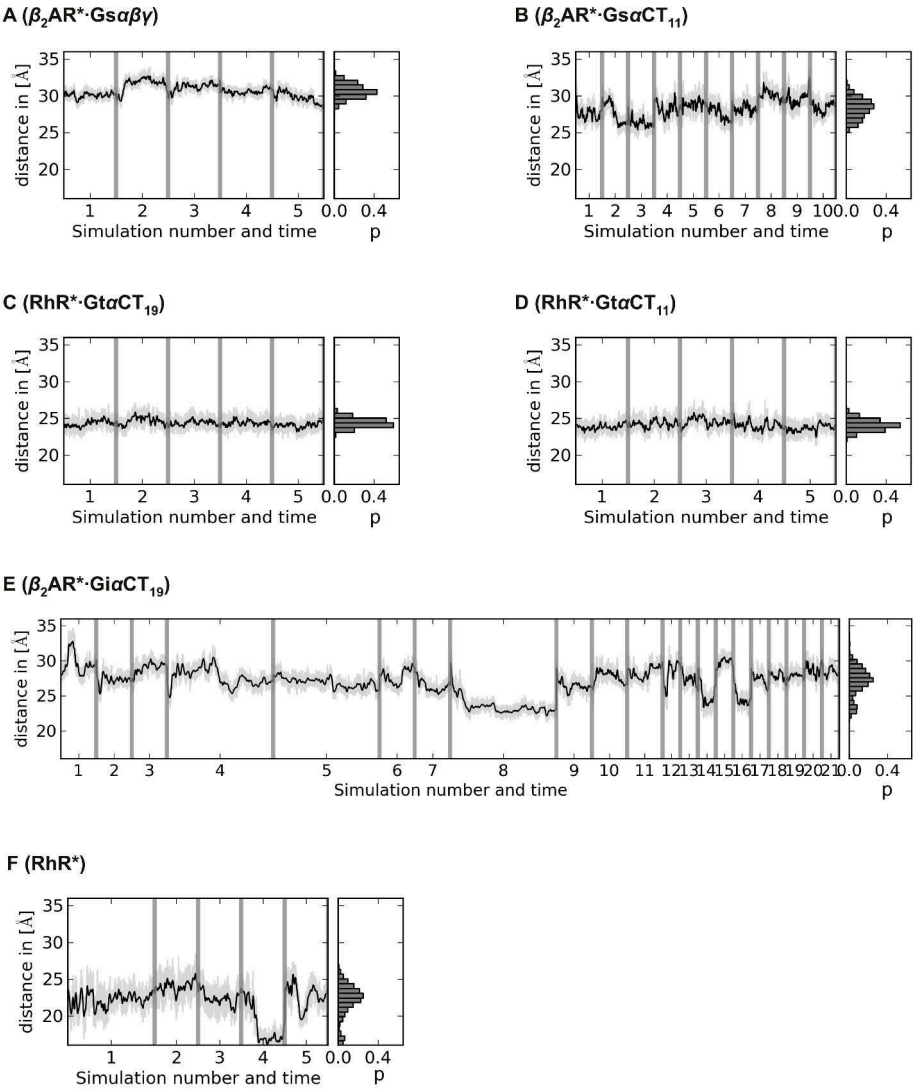

Figure A2: TM6 tilt measured as the TM2-TM6 distances of (A) β_2AR*•G$\alpha_s\beta\gamma$, (B) β_2AR*•Gα_sCT 19, (C) RhR*•Gα_tCT 19, (D) RhR*•Gα_tCT 11, (E) β_2AR*•Gα_iCT 19 and (F) RhR*. Shown are the time traces of the TM2-TM6 distances and the resulting distributions with observed probabilities p. The lengths of the individual simulations are given in Table A1. Note that N-terminal truncation to Gα_sCT 11 leads to higher observed variability of TM6 tilts compared to β_2AR*•Gα_sCT 19 or β_2AR*•G$\alpha_s\beta\gamma$, especially within individual simulations. This effect is less articulate for RhR*•Gα_tCT11 compared to Gα_tCT19. *The figure and legend are reproduced from [101]*

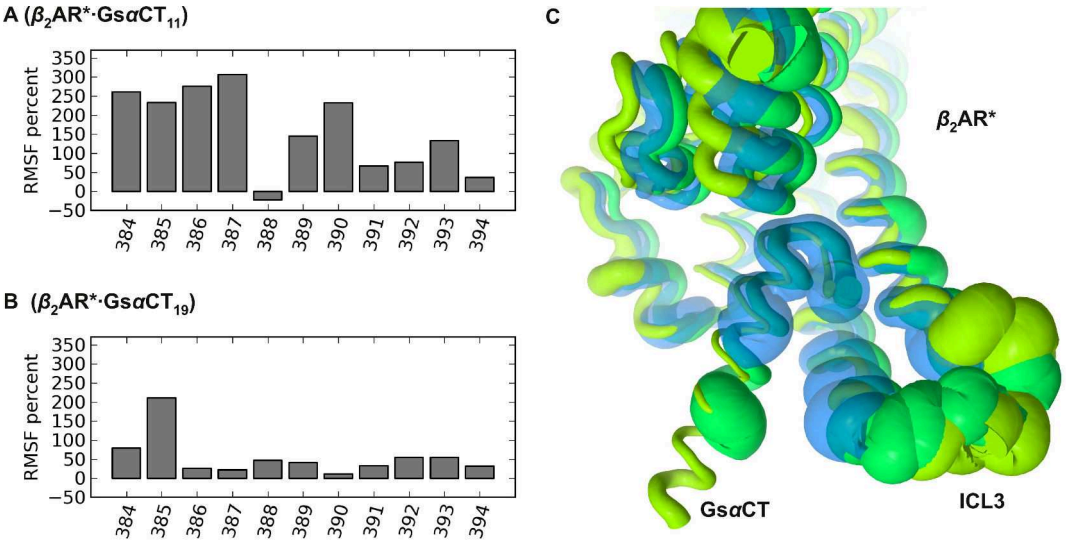

Figure A3: $G\alpha_s$CT and ICL3 flexibility of various complexes. Shown are the backbone RMSF changes of $G\alpha_s$CT relative to β_2AR*•$G\alpha_s\beta\gamma$ of (A) β_2AR*•$G\alpha_s$CT11 or (B) β_2AR*•$G\alpha_s$CT19. (C) Tube depiction of β_2AR* with $G\alpha_s$CT 11 (blue), with $G\alpha_s$CT19 (green) or with $G\alpha_s$CT from $G\alpha_s\beta\gamma$ (yellow). The thickness of the tube corresponds to the RMSF value of the respective amino acids. The RMSF is the standard deviation of atomic positions over simulation time and was calculated with the GROMACS tool g_rmsf from the MD simulations of the respective systems as listed in Table A1. Before the analysis each frame was superposed with the backbone atoms of TM1-TM7 of the initial structure. *The figure and legend are reproduced from [101]*

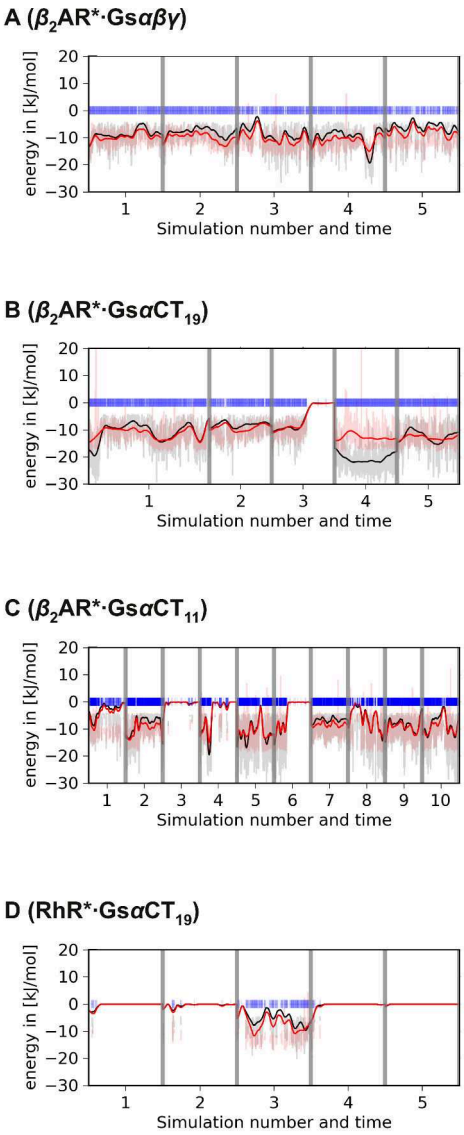

Figure A4: Cation-π interactions between β_2AR* and (A) G$\alpha_s\beta\gamma$, (B) Gα_sCT19, (C) Gα_sCT11 and between (D) RhR* and Gα_sCT19. Each plot shows the data of multiple MD simulations, each between 200 ns and 500 ns long. See Table A1 for the lengths of the individual simulations. The time series show the electrostatic (black) and hydrophobic (red, Lenard-Jones potential) contributions to the interaction energy as calculated by the CaPTURE program [160]. Blue lines denote frames where the interaction energy of cation-π interaction between $R^{3.50}$ and C391 is significant, which occurs continuously in the β_2AR* simulations with the exceptions of simulation 3 in (B) and simulations 3, 4, 6 in (C). The cation-π interaction between RhR* and Gα_sCT19 seen in simulation 3 (D) exhibits no face-to-face orientation as in β_2AR* but rather an edge-to-face orientation. *The figure and legend are reproduced from [101]*

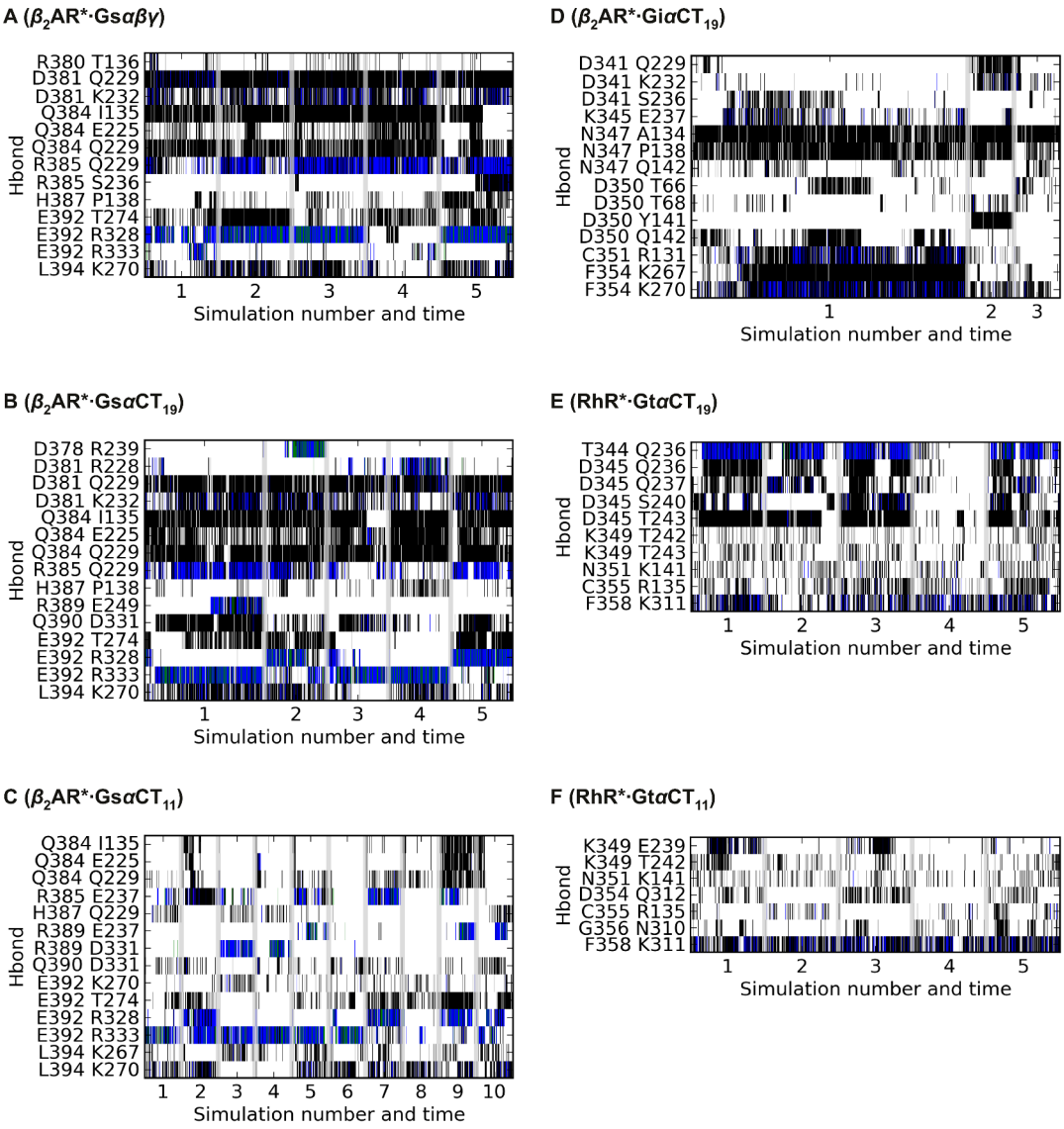

Figure A5: Polar interactions between R* and GαCT. Each plot shows the data of multiple MD simulation, each between 100 ns and 600 ns long. The lengths of the individual simulations are given in Table A1. The time series denote one (black), two (blue) or three (green) polar interactions between two residues as observed in the MD simulations of (A) β_2AR*•G$\alpha_s\beta\gamma$, (B) β_2AR*•Gα_sCT19, (C) β_2AR*•Gα_sCT 11, (D) β_2AR*•Gα_iCT19, (E) RhR*•Gα_tCT19 and (F) RhR*•Gα_tCT11. Polar interactions between GαCT and R* were calculated with the GROMACS tool g_hbond using a donor-acceptor distance cutoff at 3.6 Å and a cutoff angle of 30° for the angle given by the acceptor-donor-hydrogen atoms. In (D) only those simulations of β_2AR*•Gα_iCT19 are shown where TM6 changes its starting position and moves inward. *The figure and legend are reproduced from [101]*

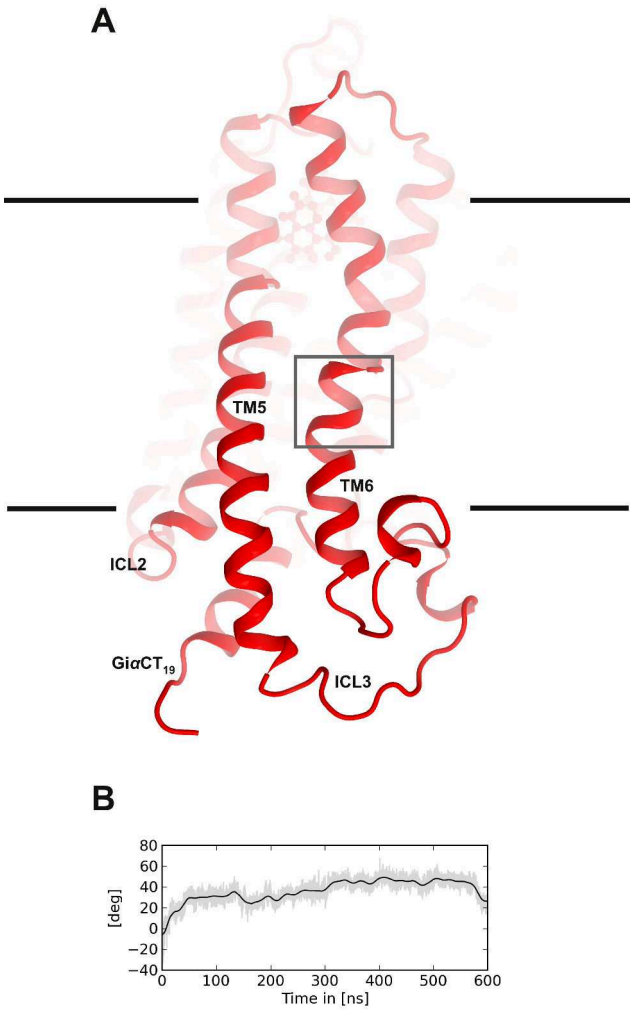

Figure A6: (A) Structuring of ICL3 that accompanies TM6 inward movement (see 15C) in the MD simulation of β_2AR*•Gα_iCT 19 used in the Umbrella Sampling. (B) Rotation of TM6 around its axis (at residues 273–280, gray box) observed in the same simulation. *The figure and legend are reproduced from [101]*

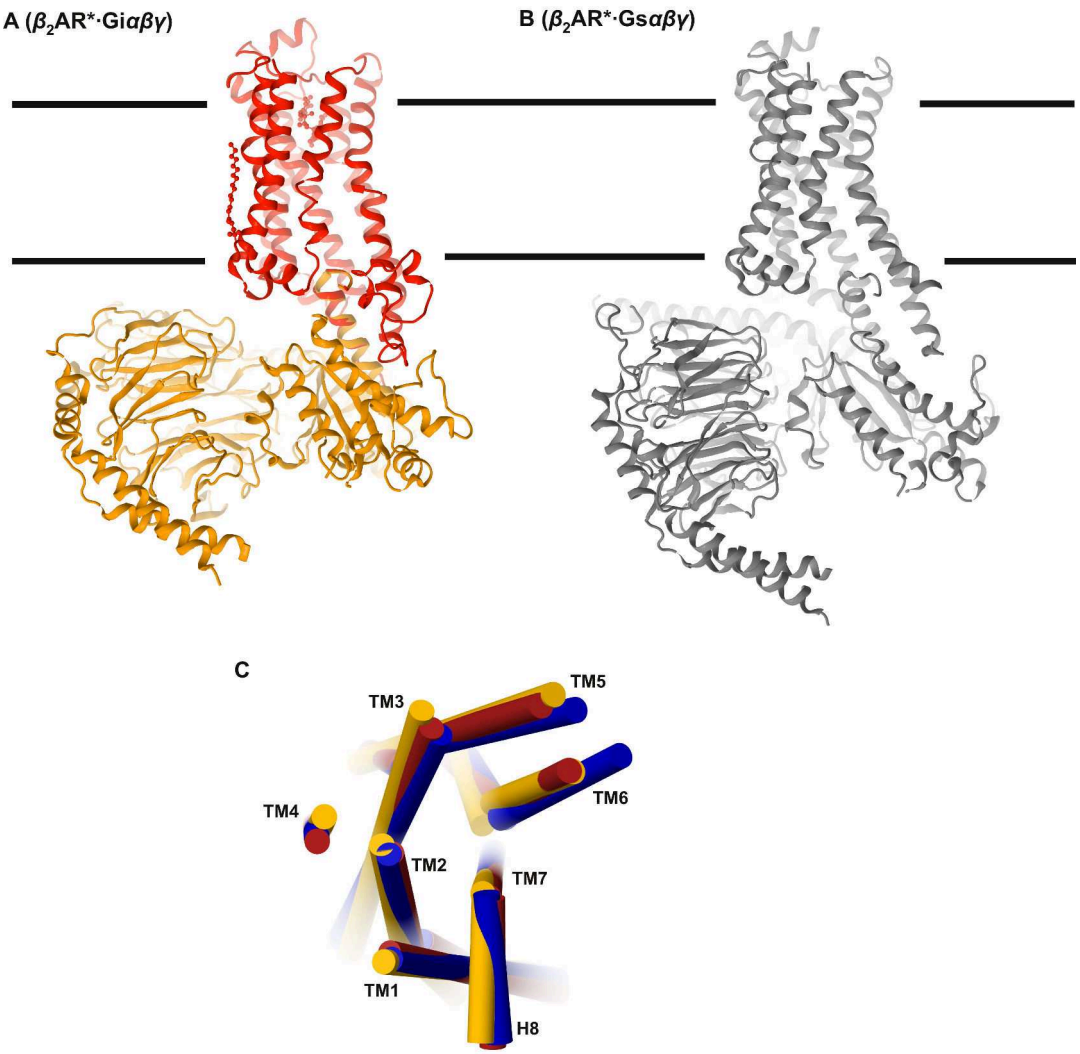

Figure A7: (A) Model of β_2AR*•G$\alpha_i\beta\gamma$ based on the arrangement of β_2AR*•Gα_iCT19 determined by classical and US MD simulation data and (B) the crystal structure of β_2AR*•G$\alpha_s\beta\gamma$ (for comparison). Apparently, there are no clashes with the membrane (black lines) nor within the complex itself. As in the β_2AR*•G$\alpha_s\beta\gamma$ complex, the arrangement in β_2AR*•G$\alpha_i\beta\gamma$ does not result in any major clashes (as shown with the inactive GDP bound G$\alpha_{i/t}$ in [78]). Comparison of both complexes, however, reveals a difference in the rotational tilt of the GαCT relative to β_2AR, resulting in a slightly different orientation of G$\alpha_i\beta\gamma$ relative to the receptor. (C) Comparison of the transmembrane helix arrangements in β_2AR*•Gα_iCT19 (red), RhR*•Gα_tCT11 (orange) and β_2AR*•G$_s$ (blue). *The figure and legend are reproduced from [101]*

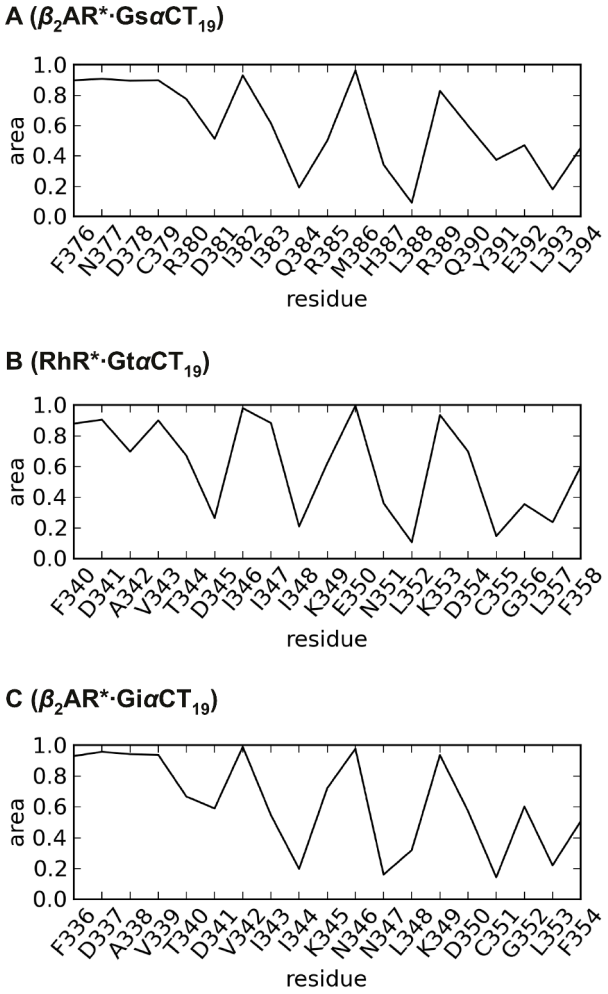

Figure A8: Per residue interaction area fraction of the $G\alpha_s$CT19 peptides observed in MD simulations of (A) β_2AR*•$G\alpha_s$CT19, (B) RhR*•$G\alpha_t$CT19 and (C) β_2AR*•$G\alpha_i$CT19. The interaction area fraction is calculated as the fraction of solvent accessible surface (SAS) (calculated with the GROMACS tool g_sas) of the R*•GαCT complex and the SAS of the free peptide. A fraction of one means a residue is completely accessible whereas completely buried residues have a fraction of zero. Residues with a fraction below 0.35 are denoted buried and those with a fraction above 0.7 are denoted accessible. In (C) only those simulations of β_2AR*•$G\alpha_i$CT 19 are shown where TM6 changes its starting position and moves inward. *The figure and legend are reproduced from [101]*

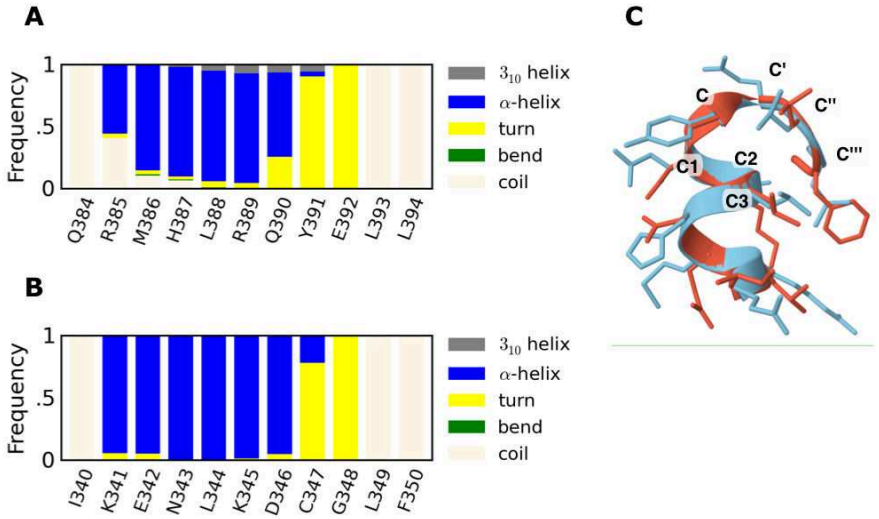

Figure A9: Secondary structure analysis of GαCT in (A) β_2AR*•Gα_sCT and (B) RhR*•Gα_tCT. Starting from the conformation of the X-ray structures from the co-crystals, analysis with the program DSSP of ten 200 ns simulations shows stable helical core regions (blue), but helix-coil transitions at the N-terminus (pale orange). The first two residues were accordingly ignored for calculation of helix axes (Figure A13) and backbone-RMSD (Figure A10). The C-terminus is defined by two residues forming a stable turn (yellow). Y391 or C347 (interacting with $R^{3.50}$ of the binding cavity) are both in position C, the so defined last residue of the helix [202]. It proceeds E392 or G348 in position C′ defined as the first turn residue. The last two C-terminal residues are assigned as coil by DSSP, but are part of the stable C-terminal reverse turn and thus were included in the analysis of backbone-RMSD. (C) Superposition of Gα_sCT and Gα_tCT. Labeling of the C-terminal reverse turn according to Aurora & Rose [202]. *The figure and legend are reproduced from [166]*

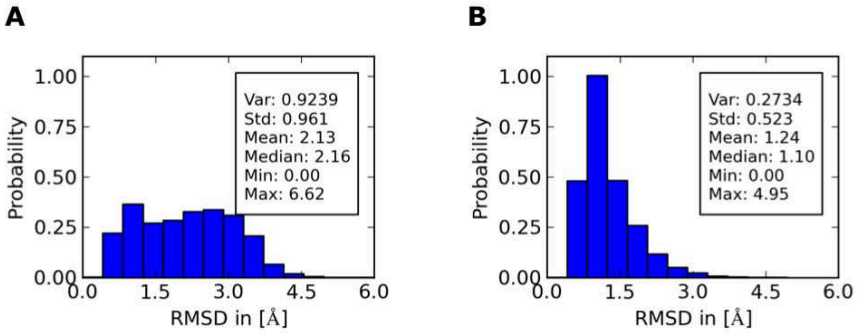

Figure A10: Mobility of (A) β_2AR*•Gα_sCT and (B) RhR*•Gα_tCT monitored by the backbone-RMSD of GαCT. The GαCT backbone-RMSD describes the movement of GαCT relative to its position in the X-ray structures. It is calculated over the complete trajectory of the simulation, after superposition to the binding cavity from the equilibrated system (also see Section 2.4.7). The histograms are calculated from ten 200 ns simulations each. For analysis only the last 9 residues were considered (for the explanation see the legend of Figure A9). *The figure and legend are reproduced from [166]*

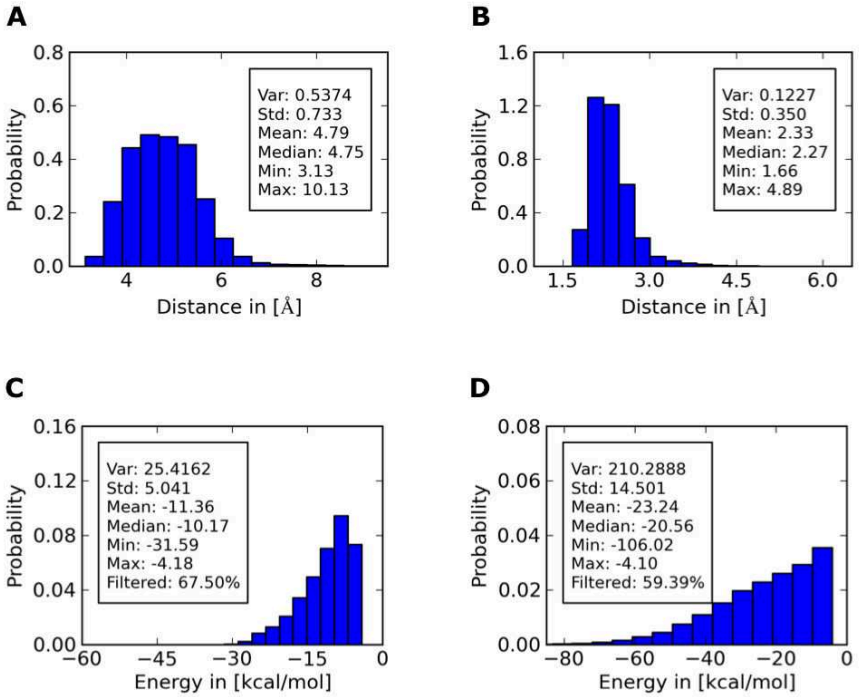

Figure A11: Distance and interaction energies with $R^{3.50}$ in β_2AR*•Gα_sCT and RhR*•Gα_tCT. Distance between (A) the center of the phenyl ring of Y391 of Gα_sCT and $R^{3.50}$ or (B) between the carbonyl oxygen of C347 of Gα_tCT and $R^{3.50}$. (C) Cation-π interaction energy between Y391 of Gα_sCT and $R^{3.50}$ or (D) hydrogen bond energy between carbonyl oxygen of C348 of Gα_tCT and $R^{3.50}$. For clarity, only energies $< -4.1\,\mathrm{kJ\,mol^{-1}}$ are shown. The histograms average over the ten 200 ns simulations, depicted in Figure A21. *The figure and legend are reproduced from [166]*

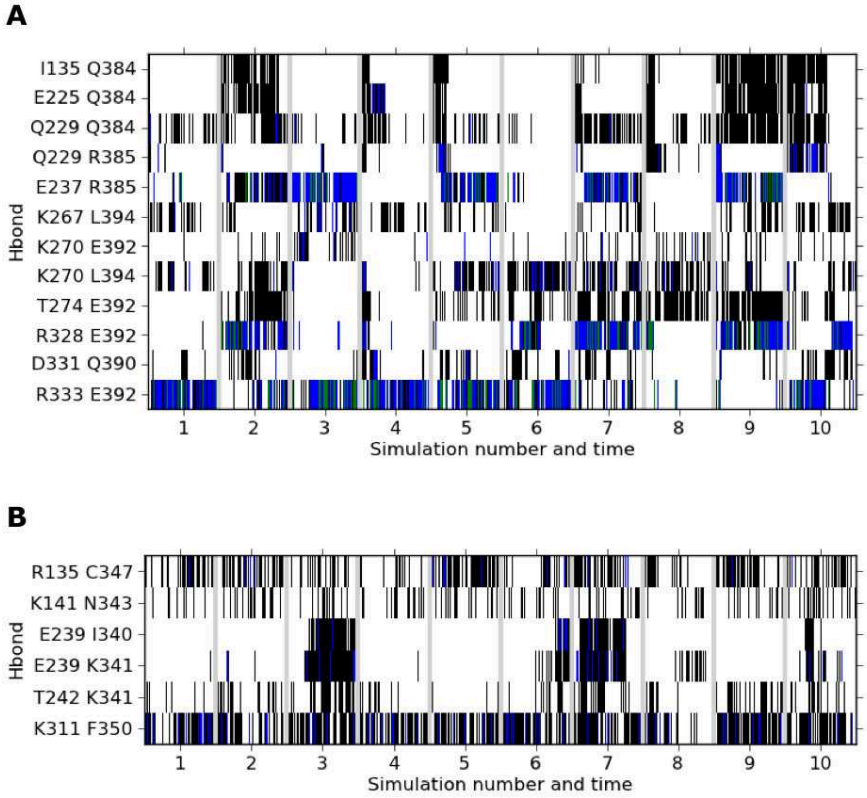

Figure A12: Potential hydrogen bonds stabilizing (A) β_2AR*•Gα_sCT and (B) RhR*•Gα_tCT. Only those interactions with a donor-acceptor distance < 3.5 Å and an angle between acceptor-donor-hydrogen $< 30°$ that persist for at least 10% of the trajectory of a simulation were considered as potential hydrogen bonds. The number of potential hydrogen bonds between two residues at a given point in time is color coded: black indicates one, blue two and green three hydrogen bonds. Gray bars separate independent 200 ns simulations. *The figure and legend are reproduced from [166]*

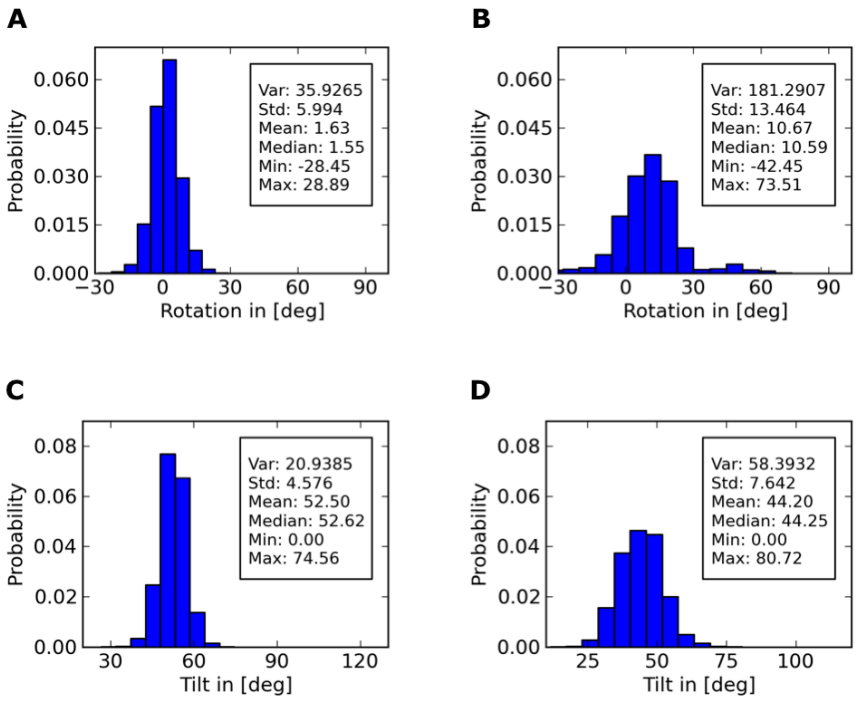

Figure A13: Mobility of GαCT in β_2AR*•Gα_sCT and RhR*•Gα_tCT measured by rotation and tilt. Histogram for the rotation around the helix axis of (A) Gα_sCT and (B) Gα_tCT. Histogram of the helix tilt motion perpendicular to the membrane plane of (C) Gα_sCT and (D) Gα_tCT. The analyses are based on 10× 200 ns simulations each (Figure A20). For definition of helix axis and tilt see Section 2.4.6. The peaks of the tilt angle distribution are clearly different, although the standard deviations overlap. *The figure and legend are reproduced from [166]*

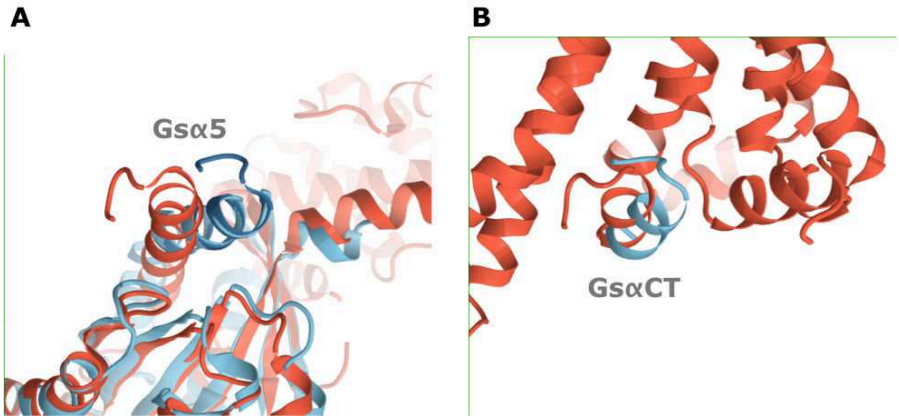

Figure A14: (A) α5 helix switch in transition from nucleotide bound G_s to nucleotide free β_2AR*•G_s[empty] and (B) rotation/ switch-like movement of $G\alpha_s$CT in the binding cavity of R*. Comparison of the α5 helix switch in G with orientations of $G\alpha_s$CT in the binding cavity of R* reveals a very similar rotation of about 70°. The nucleotide bound G_s ($G\alpha_s$[GTPγS], PDB entry 1AZT, colored in light blue) was superimposed to β_2AR*•G_s[empty] (PDB entry 3SN6, colored in red) by the Ras domain, excluding α5. The unresolved C-terminal reverse turn in $G\alpha_s$[GTPγS] was modeled from superimposing the α5 helix of β_2AR*•G_s[empty] (dark blue). The orientation of $G\alpha_s$CT in a putative R*•G[GDP] intermediate (colored in light blue) was obtained from flexible docking analysis (Figure 18, Section 2.5) and its orientation in the X-ray (colored in red) was derived from PDB entry 3SN6. Superimposition with $G\alpha_i$[GDP] (PDB entry 1GP2) instead of $G\alpha_s$[GTPγS] reveals a very similar rotation of α5. *The figure and legend are reproduced from [166]*

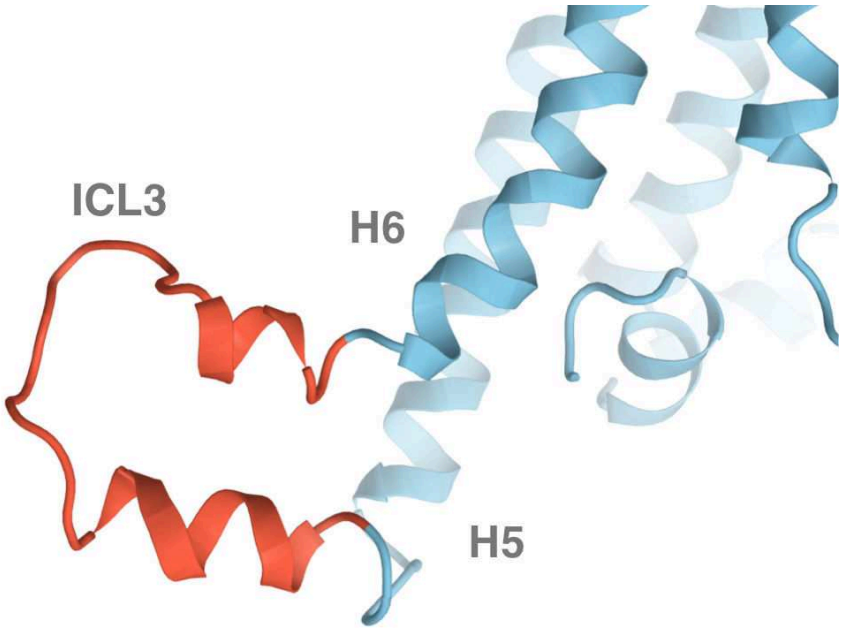

Figure A15: Structure of modeled ICL3 of β_2AR*. The protein loop was modeled with the program SuperLooper [142] which utilizes existing loops linking transmembrane helices from structures in the PDB (`http://www.rcsb.org`). See also Section 2.2 for details on the preparation of active receptor and GαCT structures and complexes. *The figure and legend are reproduced from [166]*

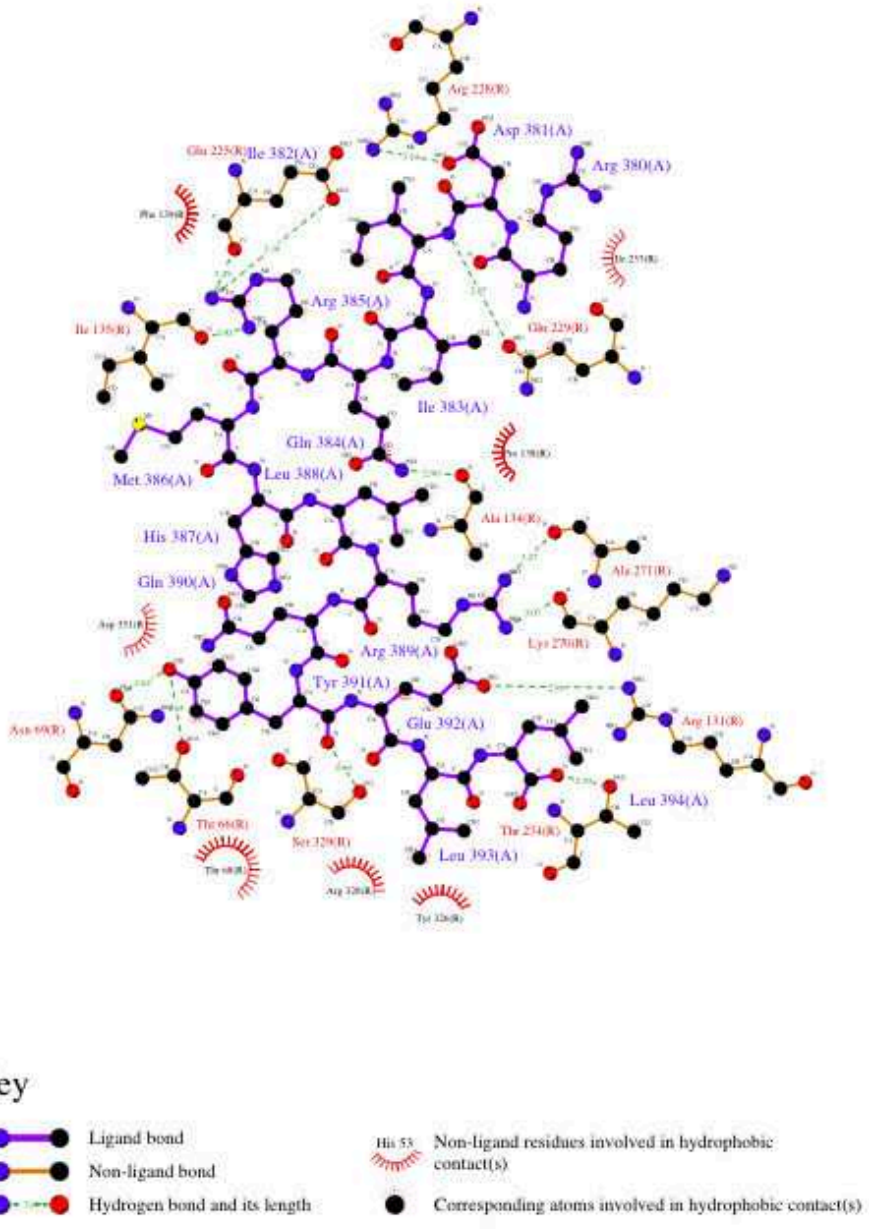

Figure A16: Interactions between β_2AR* and Gα_sCT in the putative R*• G[GDP] intermediate. Potential hydrogen bonds and van der Waals contacts were analyzed using the programs HBPLUS [203] and LIGPLOT [204]. Residues with closest distances less than 4 Å are considered to be in van der Waals contact. *The figure and legend are reproduced from [166]*

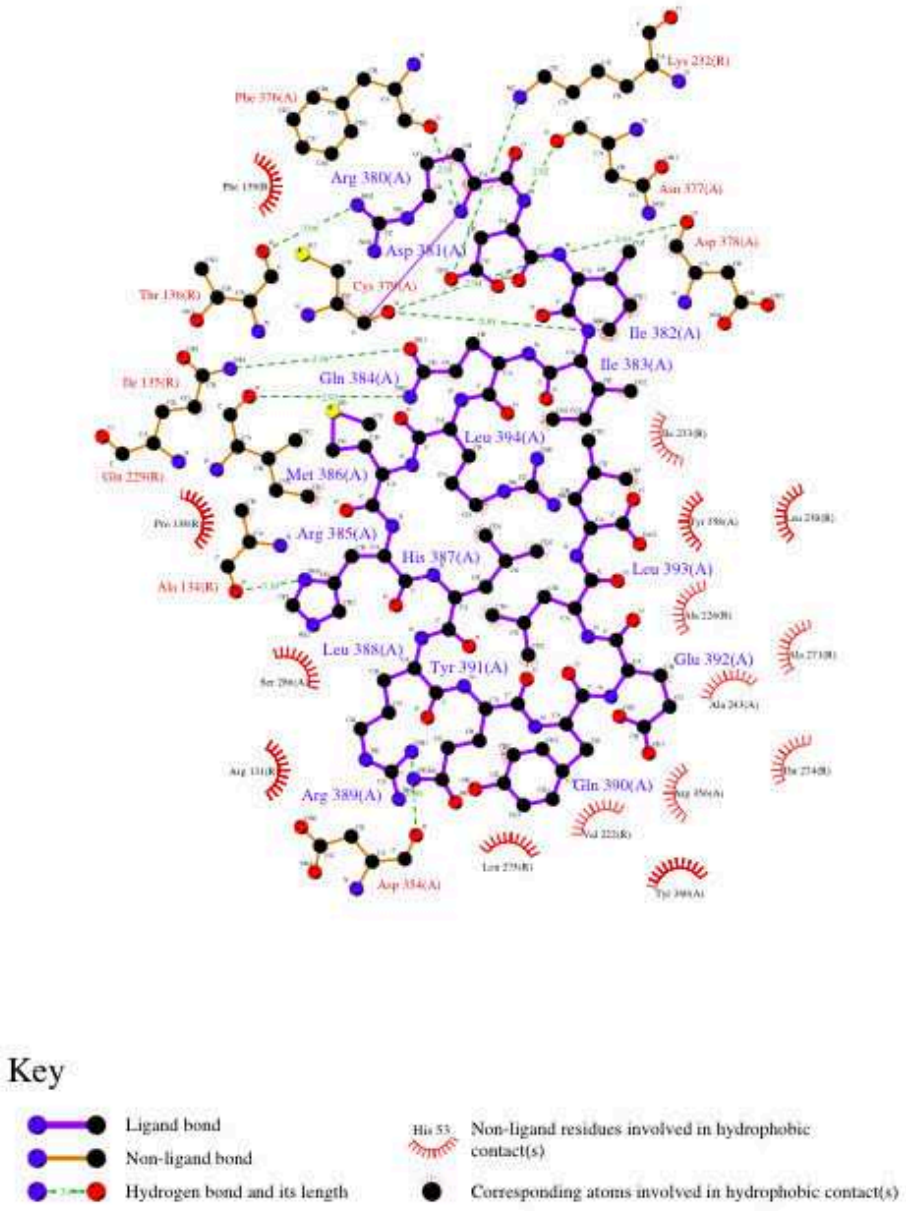

Figure A17: Interactions between β_2AR* and Gα_sCT in the position and orientation observed in the X-ray structures (PDB entry 3SN6). Potential hydrogen bonds and van der Waals contacts were analyzed using the programs HBPLUS [203] and LIGPLOT [204]. Residues with closest distances less than 4 Å are considered to be in van der Waals contact. *The figure and legend are reproduced from [166]*

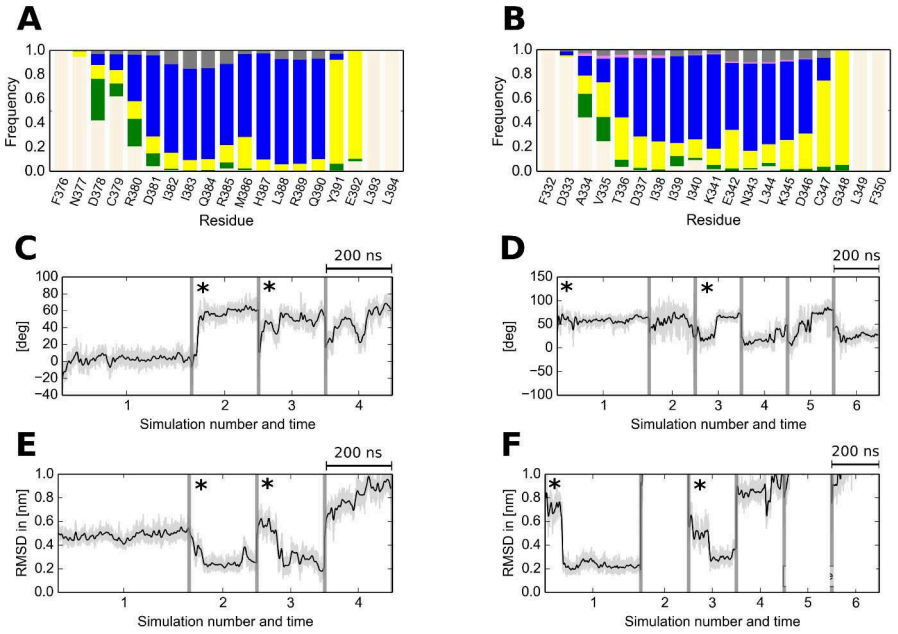

Figure A18: Secondary structure analysis of GαCT 19-mer and transitions from the intermediary to the R*•GαCT19 complex for β_2AR*•$G\alpha_s$CT19 (A, C, E) and RhR*•$G\alpha_s$CT 19 (B, D, E). (A, B) Starting from the intermediary GαCT positions obtained by flexible docking, analysis with the program DSSP of the combined/concatenated GαCT 19-mer simulations shows stable helical core regions (blue), but helix-coil transitions at the N-terminus (pale orange). See also Figure A9. The observed rotation and RMSD of GαCT are given for the individual simulations. (C, D) Change in rotation of (C) $G\alpha_s$CT or (D) $G\alpha_t$CT around its helix axis. (E, F) Backbone-RMSD of (E) $G\alpha_s$CT or (F) $G\alpha_t$CT relative to the position in the X-ray structure. In simulation marked with a *, the helix switch occurred, i. e. a rotation of about 60° and a decrease in RMSD below 4 Å was observed. *The figure and legend are reproduced from [166]*

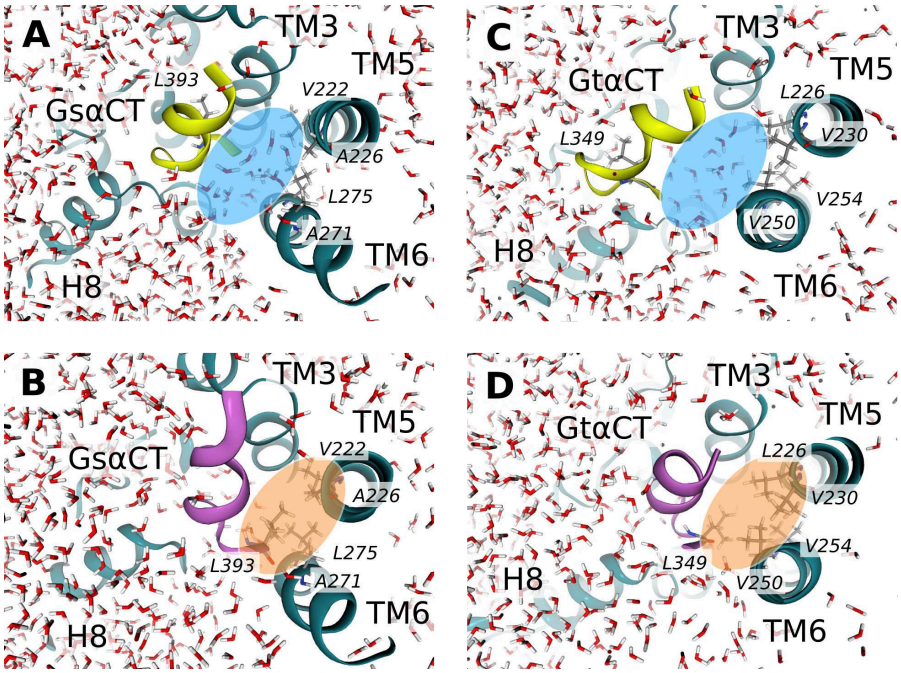

Figure A19: Water expulsion accompanying the helix switch. In the intermediate state of β_2AR*•Gα_sCT (A) and RhR*•Gα_tCT (C) a number of water molecules (blue ellipsis) are located between GαCT and TM5 and 6 of R*. During the helix switch and formation of the final β_2AR*•Gα_sCT (B) and RhR*•Gα_tCT (D) complex these water molecules are displaced in favor of hydrophobic interactions (orange ellipsis) between GαCT and TM5 and 6 of R*. For time-series data of the water expulsion and the hydrophobic patch formation see Figure A28–Figure A31. *The figure and legend are reproduced from [166]*

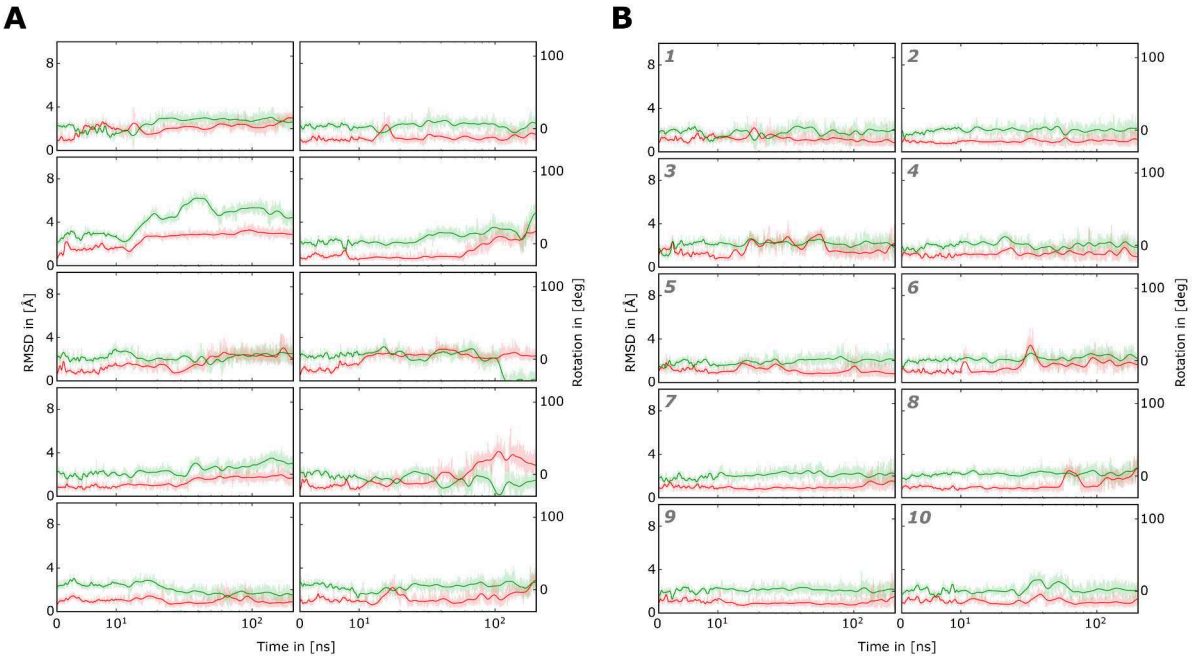

Figure A20: Mobility of GαCT in (A) β_2AR*•Gα_sCT and (B) Ops*•Gα_tCT measured by rotation and backbone RMSD. The rotation of GαCT around its helix axis during ten 200 ns simulations, starting from the GαCT position of the X-ray structure, is depicted in green. The peptide RMSD (red) describes the movement of GαCT relative to its position in the co-crystal. For definitions of helix axis, rotation and peptide RMSD see Section 2.4.6. The plots are linear for the first 10 ns and logarithmic for the remaining time. *The figure and legend are reproduced from [166]*

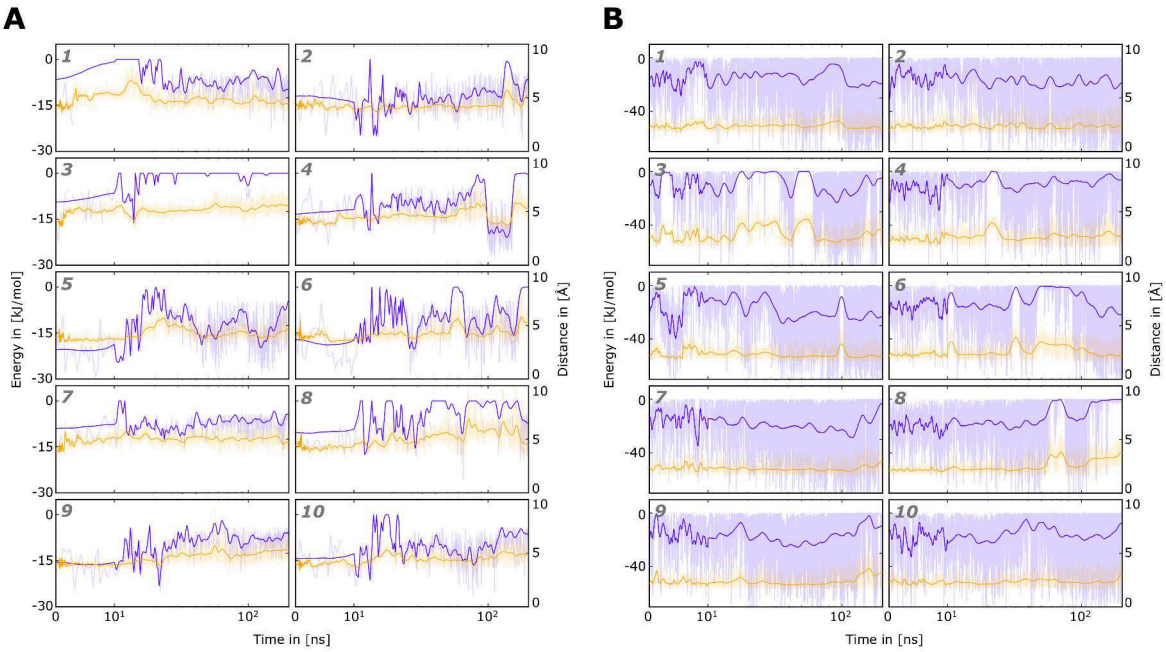

Figure A21: Distance and interaction energy of $R^{3.50}$ in β_2AR*•Gα_sCT and Ops*•Gα_tCT. (A) Distance (yellow) and cation-π interaction energy (blue) between the center of the phenylring of Y391 in Gα_sCT and $R131^{3.50}$ during ten 200 ns simulations, starting from the GαCT position of the X-ray structure. For clarity, only energies $< -4.1\,\mathrm{kJ\,mol^{-1}}$ are shown. (B) Distance (yellow) and hydrogen bond energy (blue) between the carbonyl oxygen of C348 in Gα_tCT and $R135^{3.50}$ during ten 200 ns simulations. The plots are linear for the first 10 ns and logarithmic for the remaining time. *The figure and legend are reproduced from [166]*

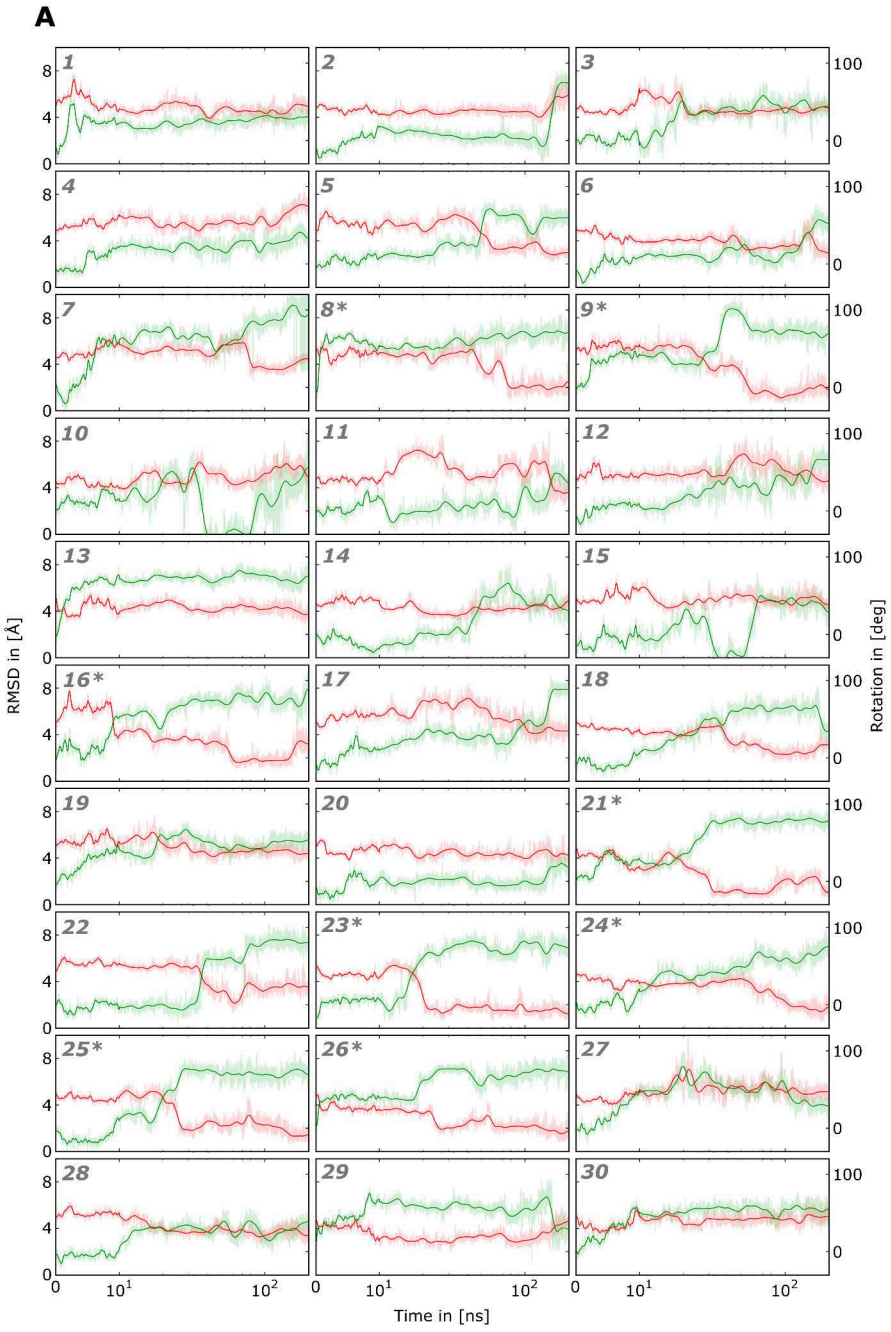

Figure A22: Mobility of $G\alpha_s CT$ in $\beta_2 AR^*$•$G\alpha_s CT$ intermediate and measured by rotation and backbone-RMSD. The rotation around the helix axis during thirty 200 ns simulations is depicted in green. The peptide RMSD (red) describes the movement of $G\alpha_s CT$ relative to its position in the crystal structure. For definitions of helix axis, rotation and peptide RMSD see Section 2.4.6. Simulations in which a switch event occurs are marked with a star (*) in the upper left corner. The plots are linear for the first 10 ns and logarithmic for the remaining time. *The figure and legend are reproduced from [166]*

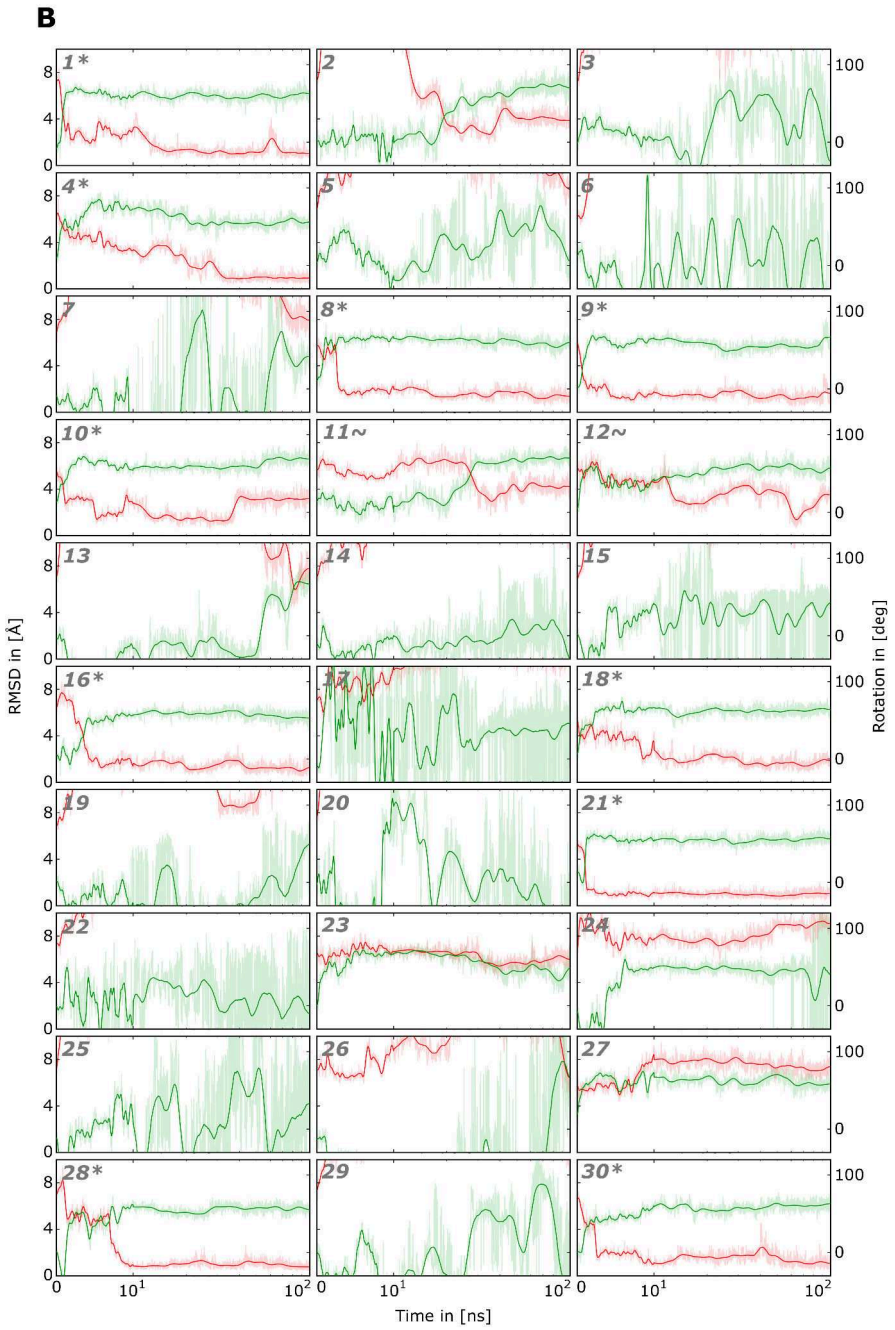

Figure A23: Mobility of $G\alpha_tCT$ in Ops*•$G\alpha_tCT$ intermediate measured by rotation and backbone-RMSD. The rotation around the helix axis during thirty 100 ns simulations is depicted in green. The peptide RMSD (red) describes the movement of $G\alpha_tCT$ relative to its position in the crystal structure. For definitions of helix axis, rotation and peptide RMSD see Section 2.4.6. Simulations in which a switch event occurs are marked with a star (*) in the upper left corner. Simulations in which $G\alpha_tCT$ adopts a stable binding mode but does not switch are marked with a tilde (~) in the upper left corner. The plots are linear for the first 10 ns (dashed gray line) and logarithmic for the remaining time. *The figure and legend are reproduced from [166]*

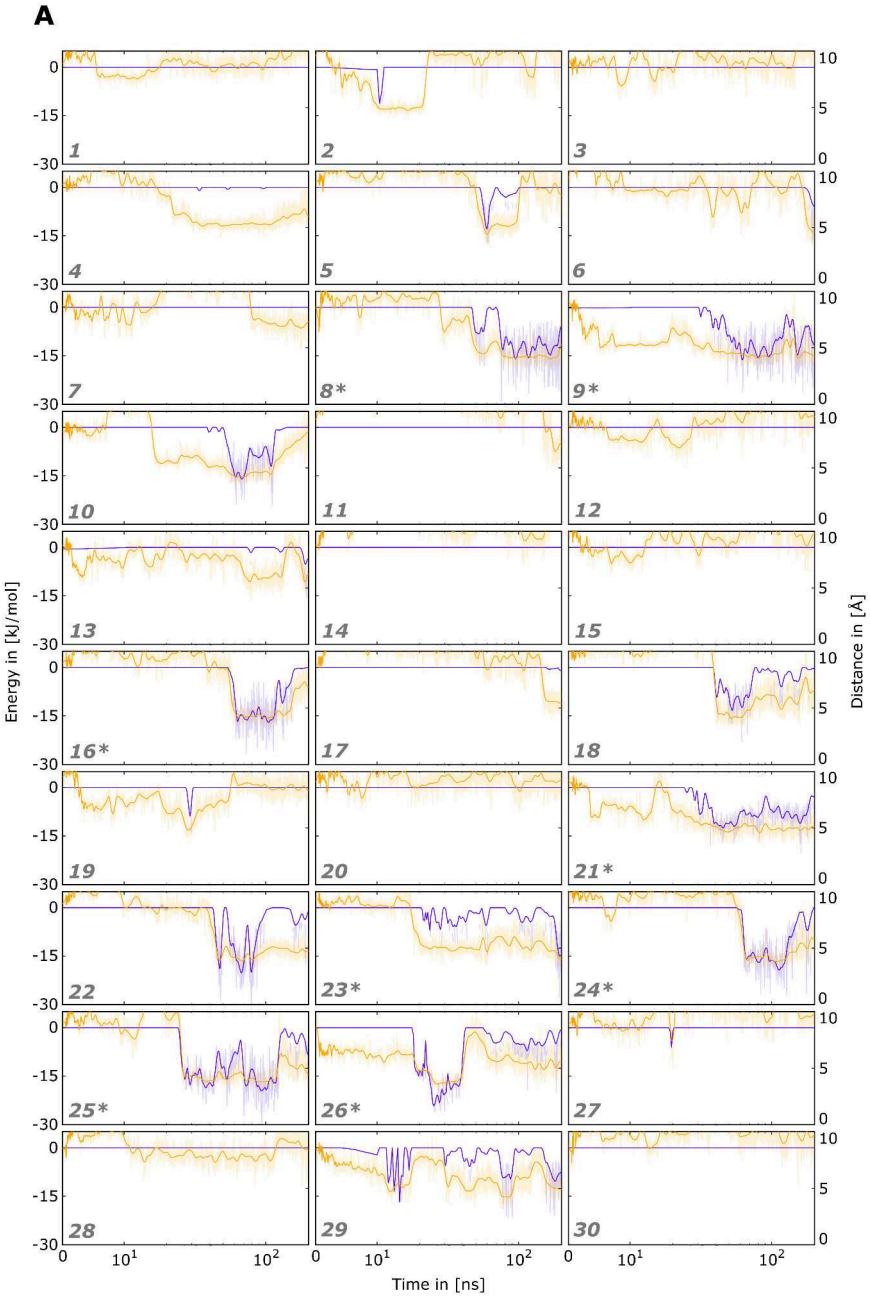

Figure A24: Distance and interaction energies of R131$^{3.50}$ in β_2AR*•Gα_sCT intermediate. Distance (yellow) and cation-π interaction energy (blue) between the center of the phenylring of Y391 in Gα_sCT and R131$^{3.50}$ during thirty 200 ns simulations. For clarity, only energies $< -4.1\,\mathrm{kJ\,mol^{-1}}$ are shown. Simulations in which a switch event occurs are marked with a star (*) in the lower left corner. The plots are linear for the first 10 ns (dashed grey line) and logarithmic for the remaining time. *The figure and legend are reproduced from [166]*

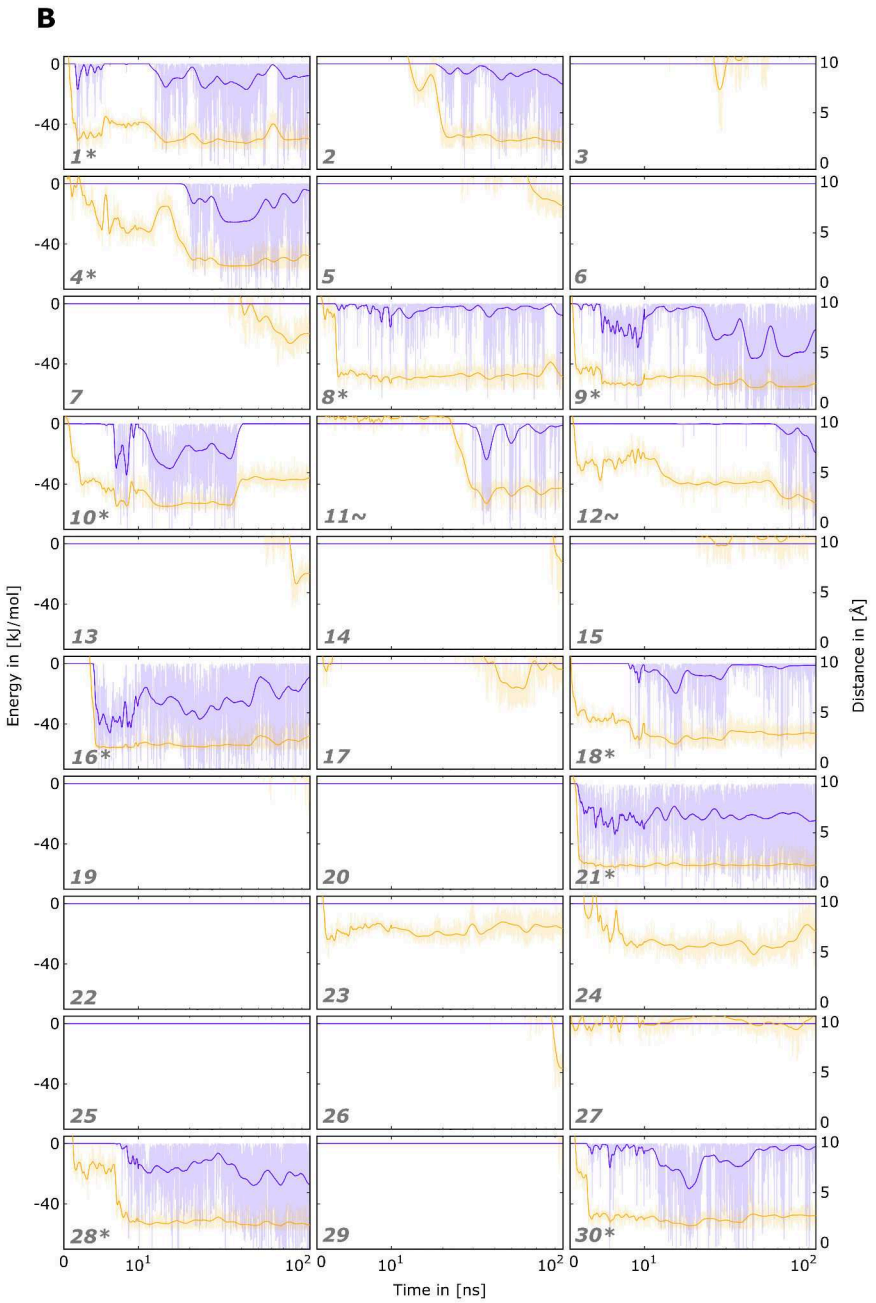

Figure A25: Distance and interaction energies of R135$^{3.50}$ Ops*•Gα_tCT intermediate. Distance (yellow) and hydrogen bond energy (blue) between carbonyl oxygen of C348 in Gα_tCT and R135$^{3.50}$ during thirty 100 ns simulations. Simulations in which a switch event occurs are marked with a star (*) in the lower left corner. Simulations in which Gα_tCT adopts a stable binding mode but does not switch are marked with a tilde (~) in the lower left corner. The plots are linear for the first 10 ns (dashed gray line) and logarithmic for the remaining time. *The figure and legend are reproduced from [166]*

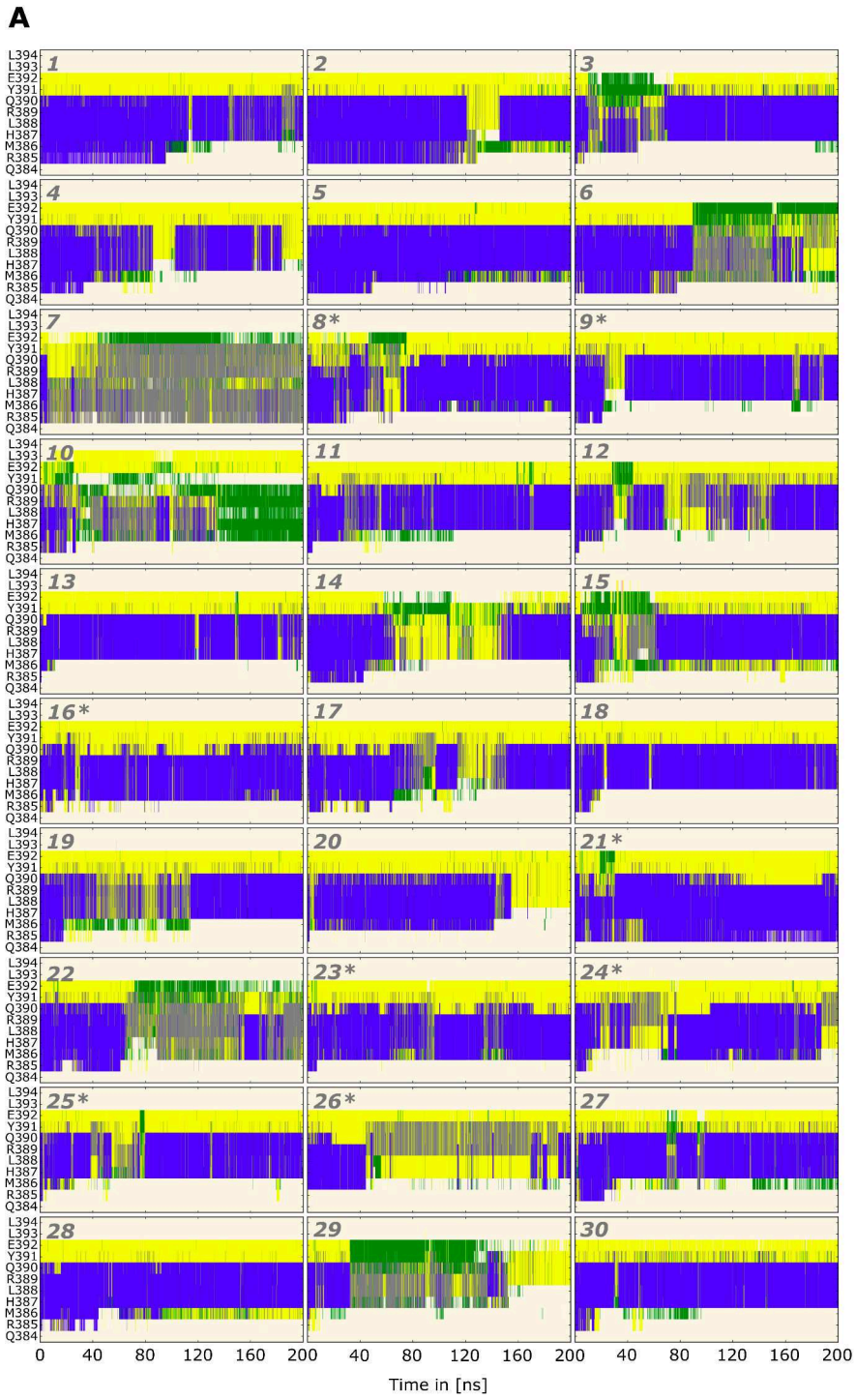

Figure A26: Secondary structure analysis of Gα_sCT in β_2AR*•Gα_sCT intermediate. For each Gα_sCT residue the secondary structure according to the program DSSP is shown over time (blue: α-helix, 3_{10}-helix: gray, turn: yellow, green: bend, coil: pale orange). Simulations in which a switch event occurs are marked with a gray star (*) in the upper left corner. *The figure and legend are reproduced from [166]*

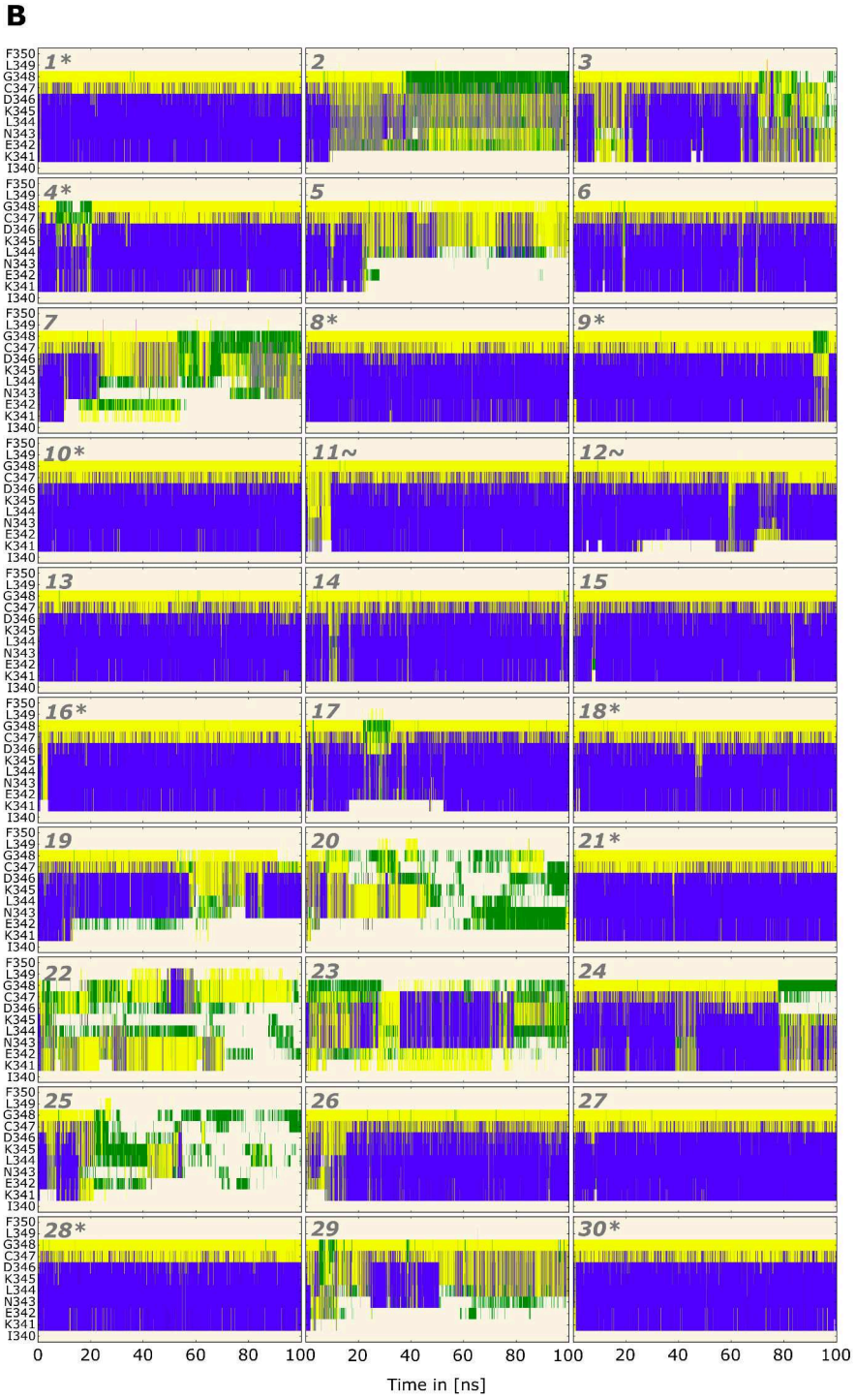

Figure A27: Secondary structure analysis of $G\alpha_t$CT in Ops*•$G\alpha_t$CT intermediate. For each $G\alpha_t$CT residue the secondary structure according to the program DSSP is shown over time (blue: α-helix, 3_{10}-helix: gray, turn: yellow, green: bend, coil: pale orange). Simulations in which a switch event occurs are marked with a gray star (*) in the upper left corner. *The figure and legend are reproduced from [166]*

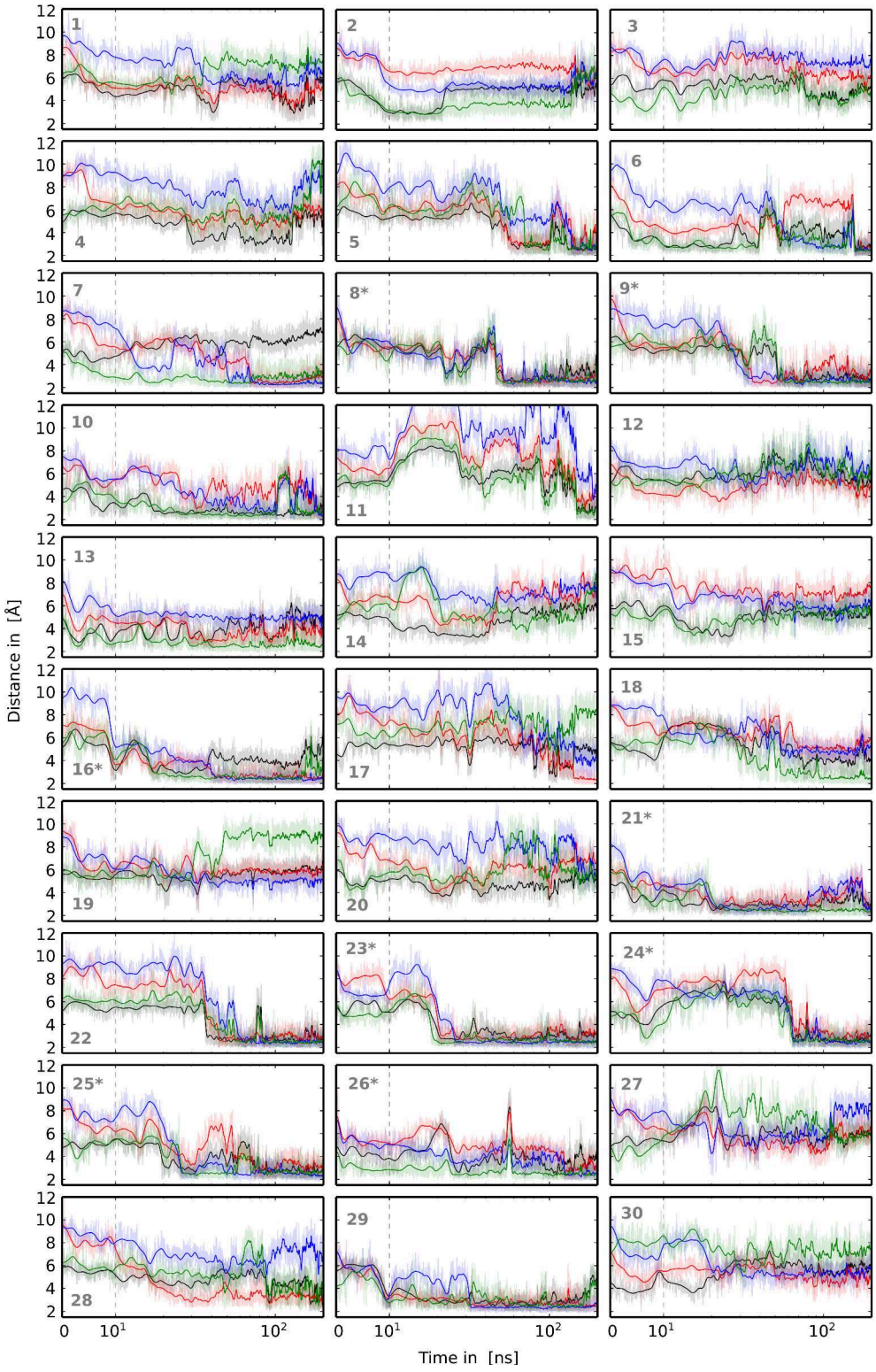

Figure A28: Hydrophobic patch between $G\alpha_s$CT and β_2AR* in β_2AR*• $G\alpha_s$CT intermediate. The minimal distances between L393 of $G\alpha_s$CT and V222 (black), A226 (red), A271 (blue) and L275 (green), respectively, of TM5 and 6 of β_2AR* are plotted over time. The plots are linear for the first 10 ns (dashed gray line) and logarithmic for the remaining time. *The figure and legend are reproduced from [166]*

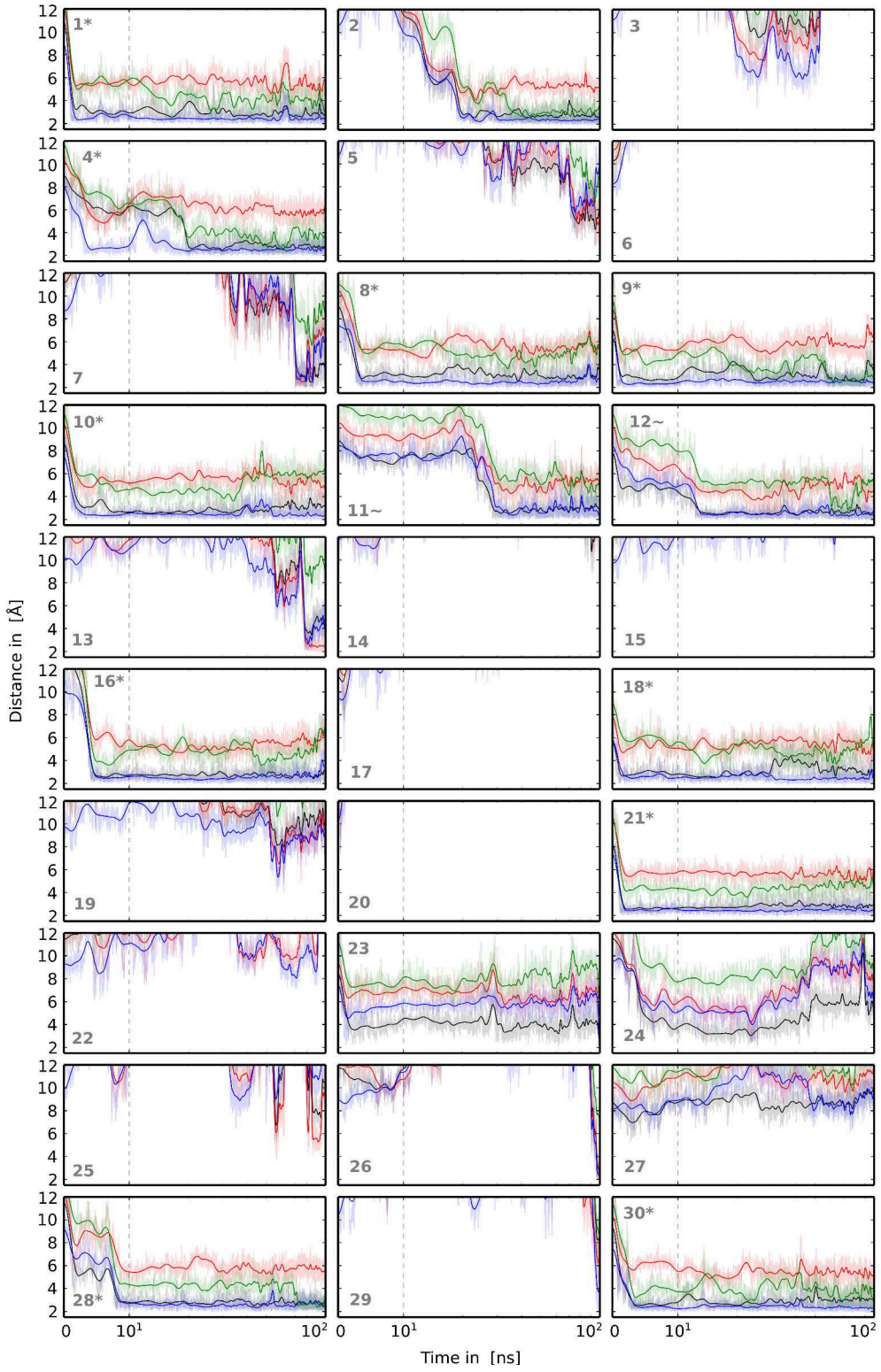

Figure A29: Hydrophobic patch between Gα_tCT and Ops* in Ops*•Gα_tCT intermediate. The minimal distances between L349 of Gα_tCT and L226 (black), V230 (red), V250 (blue) and V254 (green), respectively, of TM5 and 6 of Ops* are plotted over time. The plots are linear for the first 10 ns (dashed gray line) and logarithmic for the remaining time. *The figure and legend are reproduced from [166]*

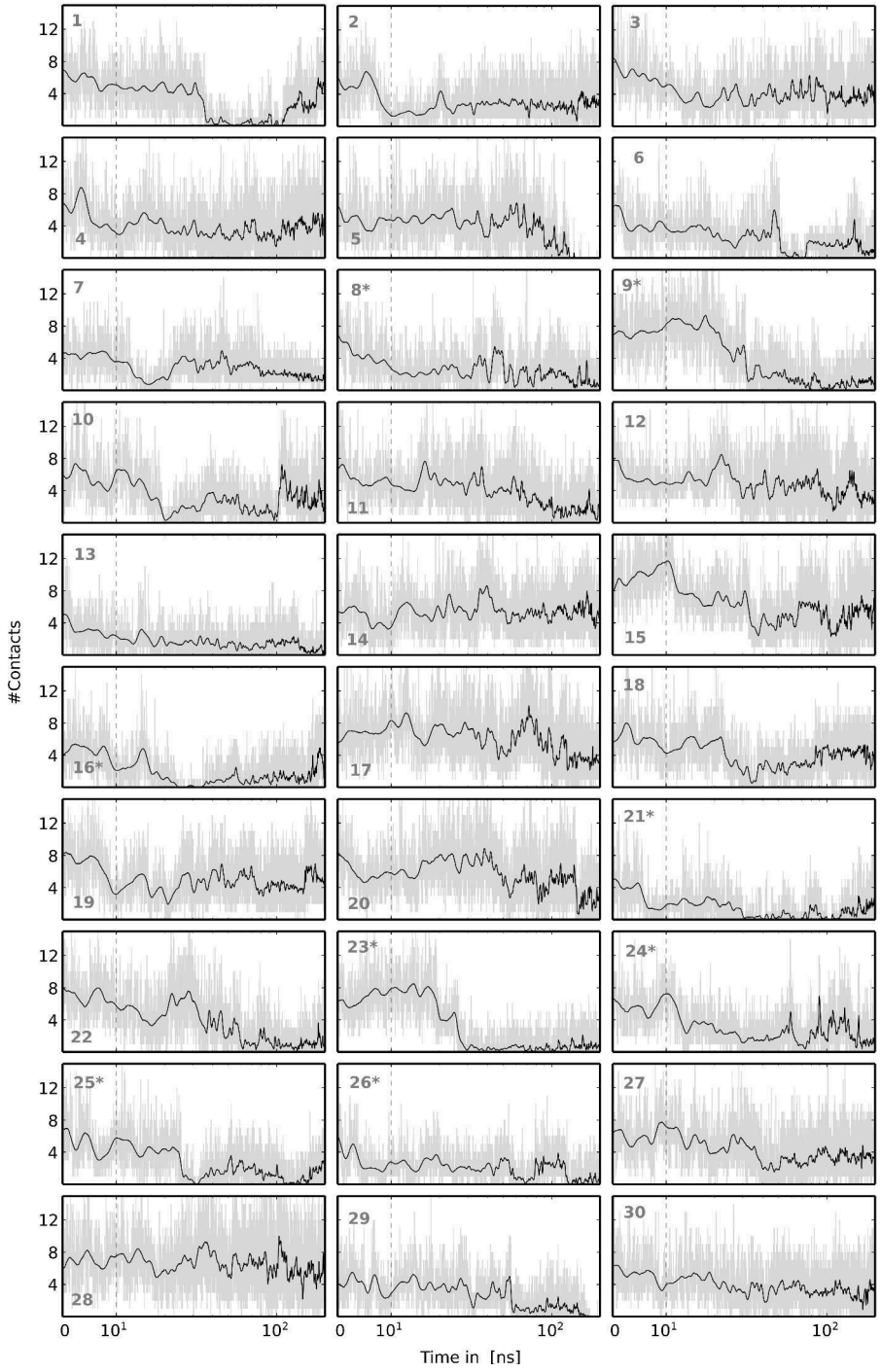

Figure A30: Hydration status of the hydrophobic patch of β_2AR* at Gα_sCT interface in β_2AR*•Gα_sCT intermediate. The number of contacts < 2.5 Å between any water molecule and V222, A226, A271 and L275 of TM5 and 6 of β_2AR* are plotted over time. Simulations in which a switch event occurs are marked with a gray star (*). The plots are linear for the first 10 ns (dashed gray line) and logarithmic for the remaining time. *The figure and legend are reproduced from [166]*

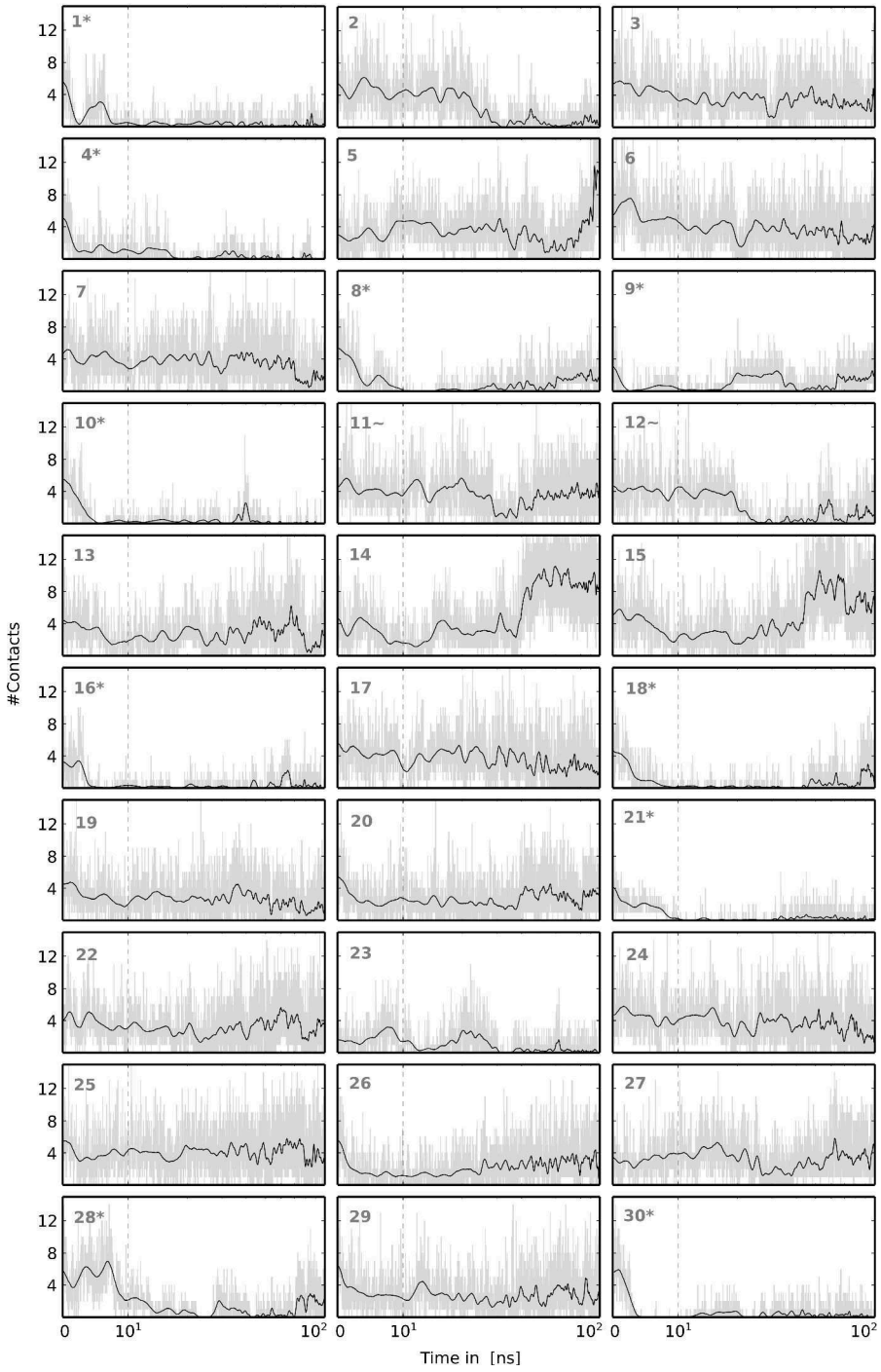

Figure A31: Hydration status of the hydrophobic patch of Ops* at $G\alpha_t CT$ interface in Ops*•$G\alpha_t CT$ intermediate. The number of contacts < 2.5 Å between any water molecule and L226, V230, V250 and V254 of TM5 and 6 of Ops* are plotted over time. Simulations in which a switch event occurs are marked with a gray star (*). The plots are linear for the first 10 ns (dashed gray line) and logarithmic for the remaining time. *The figure and legend are reproduced from [166]*

BIBLIOGRAPHY

[1] KP Hofmann, P Scheerer, PW Hildebrand, HW Choe, JH Park, M Heck, and OP Ernst. „A G protein-coupled receptor at work: the rhodopsin model.“ *Trends Biochem. Sci.* 34.11 (Nov. 2009), pp. 540–52. DOI: 10.1016/j.tibs.2009.07.005.

[2] AJ Venkatakrishnan, X Deupi, G Lebon, CG Tate, GF Schertler, and MM Babu. „Molecular signatures of G-protein-coupled receptors.“ *Nature* 494.7436 (Feb. 2013), pp. 185–94. DOI: 10.1038/nature11896.

[3] V Katritch, V Cherezov, and RC Stevens. „Structure-function of the G protein-coupled receptor superfamily.“ *Annu. Rev. Pharmacol. Toxicol.* 53 (Jan. 2013), pp. 531–56. DOI: 10.1146/annurev-pharmtox-032112-135923.

[4] RP Xiao. „Beta-adrenergic signaling in the heart: dual coupling of the beta2-adrenergic receptor to G(s) and G(i) proteins.“ *Sci. STKE* 2001.104 (Oct. 2001), re15. DOI: 10.1126/stke.2001.104.re15.

[5] HE Hamm. „The Many Faces of G Protein Signaling.“ *J. Biol. Chem.* 273.2 (Jan. 1998), pp. 669–672. DOI: 10.1074/jbc.273.2.669.

[6] SR Neves, PT Ram, and R Iyengar. „G protein pathways.“ *Science* 296.5573 (May 2002), pp. 1636–9. DOI: 10.1126/science.1071550.

[7] TG Wensel and L Stryer. „Reciprocal control of retinal rod cyclic GMP phosphodiesterase by its gamma subunit and transducin.“ *Proteins* 1.1 (Sept. 1986), pp. 90–9. DOI: 10.1002/prot.340010114.

[8] SR Sprang. „G protein mechanisms: insights from structural analysis.“ *Annu. Rev. Biochem.* 66 (Jan. 1997), pp. 639–78. DOI: 10.1146/annurev.biochem.66.1.639.

[9] U Wilden, SW Hall, and H Kühn. „Phosphodiesterase activation by photoexcited rhodopsin is quenched when rhodopsin is phosphorylated and binds the intrinsic 48-kDa protein of rod outer segments.“ *Proc. Natl. Acad. Sci. U. S. A.* 83.5 (Mar. 1986), pp. 1174–8.

[10] M Lohse, J Benovic, J Codina, M Caron, and R Lefkowitz. „beta-Arrestin: a protein that regulates beta-adrenergic receptor function.“ *Science (80-.).* 248.4962 (June 1990), pp. 1547–1550. DOI: 10.1126/science.2163110.

[11] G Venkatakrishnan and JH Exton. „Identification of Determinants in the alpha-Subunit of G(q) Required for Phospholipase C Activation.“ *J. Biol. Chem.* 271.9 (Mar. 1996), pp. 5066–5072. DOI: 10.1074/jbc.271.9.5066.

[12] JR Hepler and AG Gilman. „G proteins.“ *Trends Biochem. Sci.* 17.10 (Oct. 1992), pp. 383–7. DOI: 0968-0004(92)90005.

[13] KL Pierce, RT Premont, and RJ Lefkowitz. „Seven-transmembrane receptors.“ *Nat. Rev. Mol. Cell Biol.* 3.9 (Sept. 2002), pp. 639–50. DOI: 10.1038/nrm908.

[14] BK Shoichet and BK Kobilka. „Structure-based drug screening for G-protein-coupled receptors.“ *Trends Pharmacol. Sci.* 33.5 (May 2012), pp. 268–272. DOI: 10.1016/j.tips.2012.03.007.

[15] D Wootten, A Christopoulos, and PM Sexton. „Emerging paradigms in GPCR allostery: implications for drug discovery.“ *Nat. Rev. Drug Discov.* 12.8 (Aug. 2013), pp. 630–44. DOI: 10.1038/nrd4052.

[16] DW Corson, MC Cornwall, EF MacNichol, S Tsang, F Derguini, RK Crouch, and K Nakanishi. „Relief of opsin desensitization and prolonged excitation of rod photoreceptors by 9-desmethylretinal.“ *Proc. Natl. Acad. Sci. U. S. A.* 91.15 (July 1994), pp. 6958–62.

[17] GF Jang, V Kuksa, S Filipek, F Bartl, E Ritter, MH Gelb, KP Hofmann, and K Palczewski. „Mechanism of rhodopsin activation as examined with ring-constrained retinal analogs and the crystal structure of the ground state protein.“ *J. Biol. Chem.* 276.28 (July 2001), pp. 26148–53. DOI: 10.1074/jbc.M102212200.

[18] V Kuksa, F Bartl, T Maeda, GF Jang, E Ritter, M Heck, JP Van Hooser, Y Liang, S Filipek, MH Gelb, KP Hofmann, and K Palczewski. „Biochemical and physiological properties of rhodopsin regenerated with 11-cis-6-ring- and 7-ring-retinals.“ *J. Biol. Chem.* 277.44 (Nov. 2002), pp. 42315–24. DOI: 10.1074/jbc.M206014200.

[19] X Yao, C Parnot, X Deupi, VRP Ratnala, G Swaminath, D Farrens, and B Kobilka. „Coupling ligand structure to specific conformational switches in the beta2-adrenoceptor.“ *Nat. Chem. Biol.* 2.8 (Aug. 2006), pp. 417–22. DOI: 10.1038/nchembio801.

[20] S Terrillon and M Bouvier. „Roles of G-protein-coupled receptor dimerization.“ *EMBO Rep.* 5.1 (Jan. 2004), pp. 30–4. DOI: 10.1038/sj.embor.7400052.

[21] Y Chen-Izu, RP Xiao, LT Izu, H Cheng, M Kuschel, H Spurgeon, and EG Lakatta. „G(i)-dependent localization of beta(2)-adrenergic receptor signaling to L-type Ca(2+) channels.“ *Biophys. J.* 79.5 (Nov. 2000), pp. 2547–56. DOI: 10.1016/S0006-3495(00)76495-2.

[22] W Liu, E Chun, Aa Thompson, P Chubukov, F Xu, V Katritch, GW Han, CB Roth, LH Heitman, AP IJzerman, V Cherezov, and RC Stevens. „Structural basis for allosteric regulation of GPCRs by sodium ions.“ *Science* 337.6091 (July 2012), pp. 232–6. DOI: 10.1126/science.1219218.

[23] XC Zhang, C Cao, Y Zhou, and Y Zhao. „Proton transfer-mediated GPCR activation.“ *Protein Cell* (Oct. 2014). DOI: 10.1007/s13238-014-0106-4.

[24] O Soubias and K Gawrisch. „The role of the lipid matrix for structure and function of the GPCR rhodopsin.“ *Biochim. Biophys. Acta* 1818.2 (Feb. 2011), pp. 234–40. DOI: 10.1016/j.bbamem.2011.08.034.

[25] T Kenakin. „New concepts in pharmacological efficacy at 7TM receptors: IUPHAR review 2.“ *Br. J. Pharmacol.* 168.3 (Feb. 2013), pp. 554–75. DOI: 10.1111/j.1476-5381.2012.02223.x.

[26] JW Wisler, K Xiao, ARB Thomsen, and RJ Lefkowitz. „Recent developments in biased agonism.“ *Curr. Opin. Cell Biol.* 27 (Apr. 2014), pp. 18–24. DOI: 10.1016/j.ceb.2013.10.008.

[27] JJ Liu, R Horst, V Katritch, RC Stevens, and K Wüthrich. „Biased signaling pathways in β2-adrenergic receptor characterized by 19F-NMR.“ *Science* 335.6072 (Mar. 2012), pp. 1106–10. DOI: 10.1126/science.1215802.

[28] DL Farrens, C Altenbach, K Yang, WL Hubbell, and HG Khorana. „Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin.“ *Science (80-.).* 274.5288 (Nov. 1996), pp. 768–770.

[29] C Altenbach, AK Kusnetzow, OP Ernst, KP Hofmann, and WL Hubbell. „High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation.“ *Proc. Natl. Acad. Sci. U. S. A.* 105.21 (2008), pp. 7439–7444.

[30] JH Park, P Scheerer, KP Hofmann, HW Choe, and OP Ernst. „Crystal structure of the ligand-free G-protein-coupled receptor opsin.“ *Nature* 454.7201 (2008), pp. 183–187.

[31] P Scheerer, JH Park, PW Hildebrand, YJ Kim, N Krauss, HW Choe, KP Hofmann, and OP Ernst. „Crystal structure of opsin in its G-protein-interacting conformation.“ *Nature* 455.7212 (2008), pp. 497–502.

[32] SGF Rasmussen, BT Devree, Y Zou, AC Kruse, KY Chung, TS Kobilka, FS Thian, PS Chae, E Pardon, D Calinski, JM Mathiesen, STA Shah, JA Lyons, M Caffrey, SH Gellman, J Steyaert, G Skiniotis, WI Weis, RK Sunahara, and BK Kobilka. „Crystal structure of the β(2) adrenergic receptor-Gs protein complex.“ *Nature* 450.7168 (2011), pp. 383–387.

[33] AC Kruse, AM Ring, A Manglik, J Hu, K Hu, K Eitel, H Hübner, E Pardon, C Valant, PM Sexton, A Christopoulos, CC Felder, P Gmeiner, J Steyaert, WI Weis, KC Garcia, J Wess, and BK Kobilka. „Activation and allosteric modulation of a muscarinic acetylcholine receptor.“ en. *Nature* 504.7478 (Dec. 2013), pp. 101–6. DOI: 10.1038/nature12735.

[34] V Katritch, V Cherezov, and RC Stevens. „Diversity and modularity of G protein-coupled receptor structures.“ *Trends Pharmacol. Sci.* 33.1 (Jan. 2012), pp. 17–27. DOI: 10.1016/j.tips.2011.09.003.

[35] JA Ballesteros and H Weinstein. „Integrated Methods for the Construction of Three-Dimensional Models and Computational Probing of Structure-Function Relations in G Protein-Coupled Receptors.“ *Methods Neurosci.* 25 (1995).

[36] JM Kim. „Structure and function in rhodopsin: Rhodopsin mutants with a neutral amino acid at E134 have a partially activated conformation in the dark state.“ *Proc. Natl. Acad. Sci.* 94.26 (Dec. 1997), pp. 14273–14278.

[37] Ja Ballesteros, aD Jensen, G Liapakis, SG Rasmussen, L Shi, U Gether, and JA Javitch. „Activation of the beta 2-adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6.“ *J. Biol. Chem.* 276.31 (Aug. 2001), pp. 29171–7. DOI: 10.1074/jbc.M103747200.

[38] RO Dror, DH Arlow, DW Borhani, S Jensen Morten Øand Piana, and DE Shaw. „Identification of two distinct inactive conformations of the beta2-adrenergic receptor reconciles structural and biochemical observations.“ *Proc Natl Acad Sci U S A* 106.12 (2009), pp. 4689–4694.

[39] T Okada, M Sugihara, AN Bondar, M Elstner, P Entel, and V Buss. „The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure.“ *J. Mol. Biol.* 342.2 (2004), pp. 571–583.

[40] V Cherezov, DM Rosenbaum, MA Hanson, SGF Rasmussen, FS Thian, TS Kobilka, HJ Choi, P Kuhn, WI Weis, BK Kobilka, and RC Stevens. „High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor.“ *Science (80-.).* 318.5854 (2007), pp. 1258–1265.

[41] H Nakamichi and T Okada. „Crystallographic analysis of primary visual photochemistry.“ *Angew. Chemie (International ed English)* 45.26 (2006), pp. 4270–4273.

[42] H Nakamichi and T Okada. „Local peptide movement in the photoreaction intermediate of rhodopsin.“ *Proc. Natl. Acad. Sci. U. S. A.* 103.34 (Aug. 2006), pp. 12729–34. DOI: 10.1073/pnas.0601765103.

[43] M Elgeti, R Kazmin, M Heck, T Morizumi, E Ritter, P Scheerer, OP Ernst, F Siebert, KP Hofmann, and FJ Bartl. „Conserved Tyr223 5.58Plays Different Roles in the Activation and G-Protein Interaction of Rhodopsin.“ *J. Am. Chem. Soc.* 133.18 (2011), pp. 7159–7165.

[44] M Heck and KP Hofmann. „Maximal rate and nucleotide dependence of rhodopsin-catalyzed transducin activation: initial rate analysis based on a double displacement mechanism.“ *J. Biol. Chem.* 276.13 (Mar. 2001), pp. 10000–9. DOI: 10.1074/jbc.M009475200.

[45] PW Hildebrand, P Scheerer, JH Park, HW Choe, R Piechnick, OP Ernst, KP Hofmann, and M Heck. „A Ligand Channel through the G Protein Coupled Receptor Opsin.“ *PLoS One* 4.2 (2009), e4382.

[46] R Piechnick, E Ritter, PW Hildebrand, OP Ernst, P Scheerer, KP Hofmann, and M Heck. „Effect of channel mutations on the uptake and release of the retinal ligand in opsin.“ *Proc. Natl. Acad. Sci. U. S. A.* 109.14 (Apr. 2012), pp. 5247–52. DOI: 10.1073/pnas.1117268109.

[47] DM Rosenbaum, SGF Rasmussen, and BK Kobilka. „The structure and function of G-protein-coupled receptors.“ *Nature* 459.7245 (May 2009), pp. 356–63. DOI: 10.1038/nature08144.

[48] X Deupi and BK Kobilka. „Energy landscapes as a tool to integrate GPCR structure, dynamics, and function.“ *Physiology (Bethesda).* 25.5 (2010), pp. 293–303.

[49] X Deupi and J Standfuss. „Structural insights into agonist-induced activation of G-protein-coupled receptors.“ *Curr. Opin. Struct. Biol.* (2011), pp. 1–11.

[50] P Ghanouni, H Schambye, R Seifert, TW Lee, SGF Rasmussen, U Gether, and BK Kobilka. „The Effect of pH on beta 2 Adrenoceptor Function. EVIDENCE FOR PROTONATION-DEPENDENT ACTIVATION.“ *J. Biol. Chem.* 275.5 (Feb. 2000), pp. 3121–3127. DOI: 10.1074/jbc.275.5.3121.

[51] A Scheer, F Fanelli, T Costa, PG De Benedetti, and S Cotecchia. „Constitutively active mutants of the alpha 1B-adrenergic receptor: role of highly conserved polar amino acids in receptor activation.“ *EMBO J.* 15.14 (July 1996), pp. 3566–78.

[52] K Henzler-Wildman and D Kern. „Dynamic personalities of proteins.“ *Nature* 450.7172 (Dec. 2007), pp. 964–72. DOI: 10.1038/nature06522.

[53] S Ye, E Zaitseva, G Caltabiano, GFX Schertler, TP Sakmar, X Deupi, and R Vogel. „Tracking G-protein-coupled receptor activation using genetically encoded infrared probes.“ *Nature* 464.7293 (2010), pp. 1386–1389.

[54] RO Dror, DH Arlow, P Maragakis, TJ Mildorf, AC Pan, H Xu, DW Borhani, and DE Shaw. „Activation mechanism of the β2-adrenergic receptor.“ *Proc Natl Acad Sci U S A* 108.46 (Nov. 2011), pp. 18684–9. DOI: 10.1073/pnas.1110499108.

[55] R Nygaard, Y Zou, RO Dror, TJ Mildorf, DH Arlow, A Manglik, AC Pan, CW Liu, JJ Fung, MP Bokoch, FS Thian, TS Kobilka, DE Shaw, L Mueller, RS Prosser, and BK Kobilka. „The dynamic process of β(2)-adrenergic receptor activation.“ *Cell* 152.3 (Jan. 2013), pp. 532–42. DOI: 10.1016/j.cell.2013.01.008.

[56] M Elgeti, AS Rose, FJ Bartl, PW Hildebrand, KP Hofmann, and M Heck. „Precision vs flexibility in GPCR signaling.“ *J. Am. Chem. Soc.* 135.33 (Aug. 2013), pp. 12305–12. DOI: 10.1021/ja405133k.

[57] R Franke, B Konig, T Sakmar, H Khorana, and K Hofmann. „Rhodopsin mutants that bind but fail to activate transducin.“ *Science (80-.).* 250.4977 (Oct. 1990), pp. 123–125. DOI: 10.1126/science.2218504.

[58] OP Ernst, KP Hofmann, and TP Sakmar. „Characterization of rhodopsin mutants that bind transducin but fail to induce GTP nucleotide uptake. Classification of mutant pigments by fluorescence, nucleotide release, and flash-induced light-scattering assays.“ *J. Biol. Chem.* 270.18 (1995), pp. 10580–10586.

[59] CG Tate and GFX Schertler. „Engineering G protein-coupled receptors to facilitate their structure determination.“ *Curr. Opin. Struct. Biol.* 19.4 (Aug. 2009), pp. 386–95. DOI: 10.1016/j.sbi.2009.07.004.

[60] SGF Rasmussen, HJ Choi, JJ Fung, E Pardon, P Casarosa, PS Chae, BT Devree, DM Rosenbaum, FS Thian, TS Kobilka, A Schnapp, I Konetzki, RK Sunahara, SH Gellman, A Pautsch, J Steyaert, WI Weis, and BK Kobilka. „Structure of a nanobody-stabilized active state of the β(2) adrenoceptor.“ *Nature* 469.7329 (Jan. 2011), pp. 175–180. DOI: 10.1038/nature09648.

[61] VP Jaakola, J Prilusky, JL Sussman, and A Goldman. „G protein-coupled receptors show unusual patterns of intrinsic unfolding.“ *Protein Eng. Des. Sel.* 18.2 (Mar. 2005), pp. 103–10. DOI: 10.1093/protein/gzi004.

[62] A Venkatakrishnan, T Flock, DE Prado, ME Oates, J Gough, and M Madan Babu. „Structured and disordered facets of the GPCR fold.“ *Curr. Opin. Struct. Biol.* 27C.September (Sept. 2014), pp. 129–137. DOI: 10.1016/j.sbi.2014.08.002.

[63] PE Wright and HJ Dyson. „Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm.“ *J. Mol. Biol.* 293.2 (Oct. 1999), pp. 321–31. DOI: 10.1006/jmbi.1999.3110.

[64] DG Lambright, J Sondek, A Bohm, NP Skiba, HE Hamm, and PB Sigler. „The 2.0 A crystal structure of a heterotrimeric G protein.“ *Nature* 379.6563 (Jan. 1996), pp. 311–9. DOI: 10.1038/379311a0.

[65] RK Sunahara. „Crystal Structure of the Adenylyl Cyclase Activator Gs.“ *Science (80-.).* 278.5345 (Dec. 1997), pp. 1943–1947. DOI: 10.1126/science.278.5345.1943.

[66] HW Choe, YJ Kim, JH Park, T Morizumi, EF Pai, N Krauss, KP Hofmann, P Scheerer, and OP Ernst. „Crystal structure of metarhodopsin II.“ *Nature* 471.7340 (Mar. 2011), pp. 651–5. DOI: 10.1038/nature09789.

[67] J Standfuss, PC Edwards, A D'Antona, M Fransen, G Xie, DD Oprian, and GFX Schertler. „The structural basis of agonist-induced activation in constitutively active rhodopsin.“ *Nature* 471.7340 (Mar. 2011), pp. 656–60. DOI: 10.1038/nature09795.

[68] OG Kisselev, J Kao, JW Ponder, YC Fann, N Gautam, and GR Marshall. „Light-activated rhodopsin induces structural binding motif in G protein alpha subunit.“ *Proc. Natl. Acad. Sci. U. S. A.* 95.8 (1998), pp. 4270–4275.

[69] BW Koenig, G Kontaxis, DC Mitchell, JM Louis, BJ Litman, and A Bax. „Structure and Orientation of a G Protein Fragment in the Receptor Bound State from Residual Dipolar Couplings.“ *J. Mol. Biol.* 322.2 (Sept. 2002), pp. 441–461. DOI: 10.1016/S0022-2836(02)00745-3.

[70] BR Conklin, Z Farfel, KD Lustig, D Julius, and HR Bourne. „Substitution of three amino acids switches receptor specificity of Gq alpha to that of Gi alpha.“ *Nature* 363.6426 (May 1993), pp. 274–6.

[71] BR Conklin, P Herzmark, S Ishida, Ta Voyno-Yasenetskaya, Y Sun, Z Farfel, and HR Bourne. „Carboxyl-terminal mutations of Gq alpha and Gs alpha that alter the fidelity of receptor activation.“ *Mol. Pharmacol.* 50.4 (Oct. 1996), pp. 885–90.

[72] J Blahos, S Mary, J Perroy, C de Colle, I Brabet, J Bockaert, and JP Pin. „Extreme C terminus of G protein alpha-subunits contains a site that discriminates between Gi-coupled metabotropic glutamate receptors.“ *J. Biol. Chem.* 273.40 (Oct. 1998), pp. 25765–9.

[73] M Natochin, KG Muradov, RL McEntaffer, and NO Artemyev. „Rhodopsin Recognition by Mutant Gs Containing C-terminal Residues of Transducin.“ *J. Biol. Chem.* 275.4 (Jan. 2000), pp. 2669–2675. DOI: 10.1074/jbc.275.4.2669.

[74] R Herrmann, M Heck, P Henklein, C Kleuss, V Wray, KP Hofmann, and OP Ernst. „Rhodopsin-transducin coupling: role of the Galpha C-terminus in nucleotide exchange catalysis.“ *Vision Res.* 46.27 (Dec. 2006), pp. 4582–93. DOI: 10.1016/j.visres.2006.07.027.

[75] M Szczepek, F Beyrière, KP Hofmann, M Elgeti, R Kazmin, A Rose, FJ Bartl, D von Stetten, M Heck, ME Sommer, PW Hildebrand, and P Scheerer. „Crystal structure of a common GPCR-binding interface for G protein and arrestin.“ *Nat. Commun.* 5.May (Sept. 2014), p. 4801. DOI: 10.1038/ncomms5801.

[76] WM Oldham and HE Hamm. „Heterotrimeric G protein activation by G-protein-coupled receptors.“ *Nat. Rev. Mol. Cell Biol.* 9.1 (Jan. 2008), pp. 60–71. DOI: 10.1038/nrm2299.

[77] T Huber and TP Sakmar. „Escaping the flatlands: new approaches for studying the dynamic assembly and activation of GPCR signaling complexes.“ *Trends Pharmacol. Sci.* 32.7 (2011), pp. 410–419.

[78] P Scheerer, M Heck, A Goede, J Park, H Choe, OP Ernst, KP Hofmann, and PW Hildebrand. „Structural and kinetic modeling of an activating helix switch in the rhodopsin-transducin interface.“ *Proc Natl Acad Sci U S A* 106.26 (2009), pp. 10660–10665.

[79] NG Abdulaev, T Ngo, E Ramon, DM Brabazon, JP Marino, and KD Ridge. „The receptor-bound "empty pocket" state of the heterotrimeric G-protein alpha-subunit is conformationally dynamic.“ *Biochemistry* 45.43 (Oct. 2006), pp. 12986–97. DOI: 10.1021/bi061088h.

[80] WM Oldham, N Van Eps, AM Preininger, WL Hubbell, and HE Hamm. „Mechanism of the receptor-catalyzed activation of heterotrimeric G proteins.“ *Nat. Struct. Mol. Biol.* 13.9 (2006), pp. 772–777.

[81] N Van Eps, AM Preininger, N Alexander, AI Kaya, S Meier, J Meiler, HE Hamm, and WL Hubbell. „Interaction of a G protein with an activated receptor opens the interdomain interface in the alpha subunit.“ *Proc Natl Acad Sci U S A* 108.23 (2011), pp. 9420–9424.

[82] GH Westfield, SGF Rasmussen, M Su, S Dutta, BT Devree, KY Chung, D Calinski, G Velez-Ruiz, AN Oleskie, E Pardon, PS Chae, T Liu, S Li, VL Woods, J Steyaert, BK Kobilka, RK Sunahara, and G Skiniotis. „Structural flexibility of the G alpha s alpha-helical domain in the beta2-adrenoceptor Gs complex.“ *Proc Natl Acad Sci U S A* 108.38 (2011), pp. 16086–16091.

[83] KY Chung, SGF Rasmussen, T Liu, S Li, BT Devree, PS Chae, D Calinski, BK Kobilka, VL Woods, and RK Sunahara. „Conformational changes in the G protein Gs induced by the β2 adrenergic receptor.“ *Nature* 477.7366 (2011), pp. 611–615.

[84] T Orban, B Jastrzebska, S Gupta, B Wang, M Miyagi, MR Chance, and K Palczewski. „Conformational dynamics of activation for the pentameric complex of dimeric g protein-coupled receptor and heterotrimeric g protein.“ *Structure* 20.5 (May 2012), pp. 826–40. DOI: 10.1016/j.str.2012.03.017.

[85] JY Shim, KH Ahn, and Da Kendall. „Molecular basis of cannabinoid CB1 receptor coupling to the G protein heterotrimer Gαiβγ: identification of key CB1 contacts with the C-terminal helix α5 of Gαi.“ *J. Biol. Chem.* 288.45 (Nov. 2013), pp. 32449–65. DOI: 10.1074/jbc.M113.489153.

[86] JS Mnpotra, Z Qiao, J Cai, DL Lynch, A Grossfield, N Leioatts, DP Hurst, MC Pitman, ZH Song, and PH Reggio. „Structural Basis of G Protein-Coupled Receptor- Gi Protein Interaction: Formation of the Cannabinoid CB2 Receptor / Gi Protein Complex.“ *J. Biol. Chem.* (May 2014), pp. –31. DOI: 10.1074/jbc.M113.539916.

[87] NS Alexander, AM Preininger, AI Kaya, RA Stein, HE Hamm, and J Meiler. „Energetic analysis of the rhodopsin-G-protein complex links the α5 helix to GDP release.“ *Nat. Struct. Mol. Biol.* 21.1 (Jan. 2014), pp. 56–63. DOI: 10.1038/nsmb.2705.

[88] N Kapoor, ST Menon, R Chauhan, P Sachdev, and TP Sakmar. „Structural evidence for a sequential release mechanism for activation of heterotrimeric G proteins.“ *J. Mol. Biol.* 393.4 (Nov. 2009), pp. 882–97. DOI: 10.1016/j.jmb.2009.08.043.

[89] AI Kaya, AD Lokits, Ja Gilbert, TM Iverson, J Meiler, and HE Hamm. „A conserved phenylalanine as relay between the α5 helix and the GDP binding region of heterotrimeric Gi protein α subunit.“ *J. Biol. Chem.* 289.35 (July 2014), pp. 24475–87. DOI: 10.1074/jbc.M114.572875.

[90] TM Thaker, M Sarwar, AM Preininger, HE Hamm, and TM Iverson. „A Transient Interaction Between the P-loop and Switch I Contributes to the Allosteric Network Between Receptor and Nucleotide in Gαi1.“ *J. Biol. Chem.* 289.16 (Mar. 2014), pp. 11331–41. DOI: 10.1074/jbc.M113.539064.

[91] T Kondo, S Plaza, J Zanet, E Benrabah, P Valenti, Y Hashimoto, S Kobayashi, F Payre, and Y Kageyama. „Small peptides switch the transcriptional activity of Shavenbaby during Drosophila embryogenesis.“ *Science* 329.5989 (July 2010), pp. 336–9. DOI: 10.1126/science.1188158.

[92] A Ferrante. „For many but not for all: how the conformational flexibility of the peptide/MHCII complex shapes epitope selection.“ *Immunol. Res.* 56.1 (May 2013), pp. 85–95. DOI: 10.1007/s12026-012-8342-2.

[93] D Cox, M Brennan, and N Moran. „Integrins as therapeutic targets: lessons and opportunities.“ *Nat. Rev. Drug Discov.* 9.10 (Oct. 2010), pp. 804–20. DOI: 10.1038/nrd3266.

[94] A Manglik, AC Kruse, TS Kobilka, FS Thian, JM Mathiesen, RK Sunahara, L Pardo, WI Weis, BK Kobilka, and S Granier. „Crystal structure of the μ-opioid receptor bound to a morphinan antagonist.“ *Nature* 485.7398 (May 2012), pp. 321–6. DOI: 10.1038/nature10954.

[95] K Meindl, T Schmiederer, K Schneider, A Reicke, D Butz, S Keller, H Gühring, L Vértesy, J Wink, H Hoffmann, M Brönstrup, GM Sheldrick, and RD Süssmuth. „Labyrinthopeptins: a new class of carbacyclic lantibiotics.“ *Angew. Chem. Int. Ed. Engl.* 49.6 (Feb. 2010), pp. 1151–4. DOI: 10.1002/anie.200905773.

[96] M Ahmad, W Gu, and V Helms. „Mechanism of fast peptide recognition by SH3 domains.“ *Angew. Chem. Int. Ed. Engl.* 47.40 (Jan. 2008), pp. 7626–30. DOI: 10.1002/anie.200801856.

[97] L Sillerud and R Larson. „Design and Structure of Peptide and Peptidomimetic Antagonists of Protein- Protein Interaction.“ *Curr. Protein Pept. Sci.* 6.2 (Apr. 2005), pp. 151–169. DOI: 10.2174/1389203053545462.

[98] H Yin and AD Hamilton. „Strategies for targeting protein-protein interactions with synthetic agents.“ *Angew. Chem. Int. Ed. Engl.* 44.27 (July 2005), pp. 4130–63. DOI: 10.1002/anie.200461786.

[99] K O'Callaghan, A Kuliopulos, and L Covic. „Turning receptors on and off with intracellular pepducins: new insights into G-protein-coupled receptor drug development.“ *J. Biol. Chem.* 287.16 (Apr. 2012), pp. 12787–96. DOI: 10.1074/jbc.R112.355461.

[100] H Hamm, D Deretic, A Arendt, P Hargrave, B Koenig, and KP Hofmann. „Site of G protein binding to rhodopsin mapped with synthetic peptides from the alpha subunit.“ *Science (80-.).* 241.4867 (1988), pp. 832–835.

[101] AS Rose, M Elgeti, U Zachariae, H Grubmüller, KP Hofmann, P Scheerer, and PW Hildebrand. „Position of transmembrane helix 6 determines receptor g protein coupling specificity.“ *J. Am. Chem. Soc.* 136.32 (Aug. 2014), pp. 11244–7. DOI: 10.1021/ja5055109.

[102] a Pulvermüller, K Schroder, T Fischer, and KP Hofmann. „Interactions of metarhodopsin II. Arrestin peptides compete with arrestin and transducin.“ *J. Biol. Chem.* 275.48 (Dec. 2000), pp. 37679–85. DOI: 10.1074/jbc.M006776200.

[103] P Csermely, R Palotai, and R Nussinov. „Induced fit, conformational selection and independent dynamic segments: an extended view of binding events.“ *Trends Biochem. Sci.* 35.10 (Oct. 2010), pp. 539–546. DOI: 10.1016/j.tibs.2010.04.009.

[104] MB Prigozhin and M Gruebele. „Microsecond folding experiments and simulations: a match is made.“ *Phys. Chem. Chem. Phys.* 15.10 (Mar. 2013), pp. 3372–88. DOI: 10.1039/c3cp43992e.

[105] M Gruebele. „The fast protein folding problem.“ *Annu. Rev. Phys. Chem.* 50 (Jan. 1999), pp. 485–516. DOI: 10.1146/annurev.physchem.50.1.485.

[106] K Lindorff-Larsen, S Piana, RO Dror, and DE Shaw. „How fast-folding proteins fold.“ *Science* 334.6055 (Oct. 2011), pp. 517–20.

[107] JA McCammon, BR Gelin, and M Karplus. „Dynamics of folded proteins.“ *Nature* 267.5612 (June 1977), pp. 585–590. DOI: 10.1038/267585a0.

[108] J Gumbart, Y Wang, A Aksimentiev, E Tajkhorshid, and K Schulten. „Molecular dynamics simulations of proteins in lipid bilayers.“ *Curr. Opin. Struct. Biol.* 15.4 (2005), pp. 423–431.

[109] C Kutzner, H Grubmüller, BL de Groot, and U Zachariae. „Computational electrophysiology: the molecular dynamics of ion channel permeation and selectivity in atomistic detail.“ *Biophys. J.* 101.4 (Aug. 2011), pp. 809–17. DOI: 10.1016/j.bpj.2011.06.010.

[110] JL Klepeis, K Lindorff-Larsen, RO Dror, and DE Shaw. „Long-timescale molecular dynamics simulations of protein structure and function.“ *Curr. Opin. Struct. Biol.* 19.2 (Apr. 2009), pp. 120–7. DOI: 10.1016/j.sbi.2009.03.004.

[111] J Chen, W Im, and CL Brooks. „Refinement of NMR structures using implicit solvent and advanced sampling techniques.“ *J. Am. Chem. Soc.* 126.49 (Dec. 2004), pp. 16038–47. DOI: 10.1021/ja047624f.

[112] Z Gattin, J Schwartz, RI Mathad, B Jaun, and WF van Gunsteren. „Interpreting experimental data by using molecular simulation instead of model building.“ *Chemistry* 15.26 (June 2009), pp. 6389–98. DOI: 10.1002/chem.200802523.

[113] D Provasi and M Filizola. „Putative active states of a prototypic g-protein-coupled receptor from biased molecular dynamics.“ *Biophys. J.* 98.10 (2010), pp. 2347–2355.

[114] D Provasi, MC Artacho, A Negri, JC Mobarec, and M Filizola. „Ligand-Induced Modulation of the Free-Energy Landscape of G Protein-Coupled Receptors Explored by Adaptive Biasing Techniques.“ *PLoS Comput. Biol.* 7.10 (2011). Ed. by R Nussinov, e1002193.

[115] Y Miao, SE Nichols, PM Gasper, VT Metzger, and JA McCammon. „Activation and dynamic network of the M2 muscarinic receptor.“ *Proc. Natl. Acad. Sci. U. S. A.* 110.27 (July 2013), pp. 10982–7. DOI: 10.1073/pnas.1309755110.

[116] KJ Kohlhoff, D Shukla, M Lawrenz, GR Bowman, DE Konerding, D Belov, RB Altman, and VS Pande. „Cloud-based simulations on Google Exacycle reveal ligand modulation of GPCR activation pathways.“ *Nat. Chem.* 6.1 (Jan. 2014), pp. 15–21. DOI: 10.1038/nchem.1821.

[117] N Leioatts, P Suresh, TD Romo, and A Grossfield. „Structure-Based Simulations Reveal Concerted Dynamics of GPCR Activation.“ *Proteins* (May 2014). DOI: 10.1002/prot.24617.

[118] MC Zwier and LT Chong. „Reaching biological timescales with all-atom molecular dynamics simulations.“ *Curr. Opin. Pharmacol.* 10.6 (Dec. 2010), pp. 745–52. DOI: 10.1016/j.coph.2010.09.008.

[119] LM Simpson, B Taddese, ID Wall, and CA Reynolds. „Bioinformatics and molecular modelling approaches to GPCR oligomerization.“ *Curr. Opin. Pharmacol.* 10.1 (2010), pp. 30–37.

[120] Y Polyhach, E Bordignon, and G Jeschke. „Rotamer libraries of spin labelled cysteines for protein studies.“ *Phys. Chem. Chem. Phys.* 13.6 (Feb. 2011), pp. 2356–2366. DOI: 10.1039/c0cp01865a.

[121] R Fine, G Dimmler, and C Levinthal. „FASTRUN: a special purpose, hardwired computer for molecular simulation.“ *Proteins* 11.4 (Jan. 1991), pp. 242–53. DOI: 10.1002/prot.340110403.

[122] T Narumi, M Taiji, M Ikei, Y Ohno, N Okimoto, T Koishi, A Suenaga, N Futatsugi, R Yanai, R Himeno, and S Fujikawa. „A 55 TFLOPS simulation of amyloid-forming peptides from yeast prion Sup35 with the special-purpose computer system MDGRAPE-3.“ In: *Proc. 2006 ACM/IEEE Conf. Supercomput. - SC '06*. New York, New York, USA: ACM Press, 2006, p. 49. DOI: 10.1145/1188455.1188506.

[123] DE Shaw, KJ Bowers, E Chow, MP Eastwood, DJ Ierardi, JL Klepeis, JS Kuskin, RH Larson, K Lindorff-Larsen, P Maragakis, MA Moraes, RO Dror, S Piana, Y Shan, B Towles, JK Salmon, JP Grossman, KM Mackenzie, JA Bank, C Young, MM Deneroff, and B Batson. „Millisecond-scale molecular dynamics simulations on Anton.“ In: *Proc. Conf. High Perform. Comput. Networking, Storage Anal. - SC '09*. c. New York, New York, USA: ACM Press, 2009, p. 1. DOI: 10.1145/1654059.1654099.

[124] DE Shaw, P Maragakis, K Lindorff-Larsen, S Piana, RO Dror, MP Eastwood, Ja Bank, JM Jumper, JK Salmon, Y Shan, and W Wriggers. „Atomic-level characterization of the structural dynamics of proteins.“ *Science* 330.6002 (Oct. 2010), pp. 341–6. DOI: 10.1126/science.1187409.

[125] RO Dror, HF Green, C Valant, DW Borhani, JR Valcourt, AC Pan, DH Arlow, M Canals, JR Lane, R Rahmani, JB Baell, PM Sexton, A Christopoulos, and DE Shaw. „Structural basis for modulation of a G-protein-coupled receptor by allosteric drugs.“ *Nature* 503.7475 (Nov. 2013), pp. 295–9. DOI: 10.1038/nature12595.

[126] B Hess, H Bekker, HJC Berendsen, and JGEM Fraaije. „LINCS: A linear constraint solver for molecular simulations.“ *J. Comput. Chem.* 18.12 (Sept. 1997), pp. 1463–1472. DOI: 10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

[127] S Miyamoto and PA Kollman. „SETTLE: an analytical version of the SHAKE and RATTLE algorithm for rigid water models.“ *J. Comput. Chem.* 13.8 (1992), pp. 952–962.

[128] DN White. „A computationally efficient alternative to the Buckingham potential for molecular mechanics calculations.“ *J. Comput. Aided. Mol. Des.* 11.5 (Sept. 1997), pp. 517–21. DOI: 10.1023/A:1007911511862.

[129] K Lindorff-Larsen, S Piana, K Palmo, P Maragakis, JL Klepeis, RO Dror, and DE Shaw. „Improved side-chain torsion potentials for the Amber ff99SB protein force field.“ *Proteins* 78.8 (2010), pp. 1950–1958.

[130] JE Stone, DJ Hardy, IS Ufimtsev, and K Schulten. „GPU-accelerated molecular modeling coming of age.“ *J. Mol. Graph. Model.* 29.2 (Sept. 2010), pp. 116–25. DOI: 10.1016/j.jmgm.2010.06.010.

[131] DJ Hardy, JE Stone, and K Schulten. „Multilevel Summation of Electrostatic Potentials Using Graphics Processing Units.“ *Parallel Comput.* 35.3 (Mar. 2009), pp. 164–177. DOI: 10.1016/j.parco.2008.12.005.

[132] Y Andoh, N Yoshii, K Fujimoto, K Mizutani, H Kojima, A Yamada, S Okazaki, K Kawaguchi, H Nagao, K Iwahashi, F Mizutani, K Minami, Si Ichikawa, H Komatsu, S Ishizuki, Y Takeda, and M Fukushima. „MODYLAS: A Highly Parallelized General-Purpose Molecular Dynamics Simulation Program for Large-Scale Systems with Long-Range Forces Calculated by Fast Multipole Method (FMM) and Highly Scalable Fine-Grained New Parallel Processing Algorithms.“ *J. Chem. Theory Comput.* 9.7 (July 2013), pp. 3201–3209. DOI: 10.1021/ct400203a.

[133] L Zhang and J Hermans. „Hydrophilicity of cavities in proteins.“ *Proteins* 24.4 (Apr. 1996), pp. 433–8. DOI: 10.1002/(SICI)1097-0134(199604)24:4<433::AID-PROT3>3.0.CO;2-F.

[134] K Fahmy, F Jager, M Beck, TA Zvyaga, TP Sakmar, and F Siebert. „Protonation states of membrane-embedded carboxylic acid groups in rhodopsin and metarhodopsin II : A Fourier-transform infrared spectroscopy study of site-directed mutants.“ *Proc. Natl. Acad. Sci. USA* 90.November (1993), pp. 10206–10210.

[135] P Rath, L DeCaluwe, and P Bovee-Geurts. „Fourier transform infrared difference spectroscopy of rhodopsin mutants: light activation of rhodopsin causes hydrogen-bonding change in residue aspartic acid-83.“ *Biochemistry* 32.39 (1993).

[136] F Jaeger, K Fahmy, and T Sakmar. „Identification of glutamic acid 113 as the Schiff base proton acceptor in the metarhodopsin II photointermediate of rhodopsin.“ *Biochemistry* 33 (1994), pp. 10878–10882.

[137] K Fahmy, T Sakmar, and F Siebert. „Transducin-dependent protonation of glutamic acid 134 in rhodopsin.“ *Biochemistry* 39 (2000), pp. 10607–10612.

[138] J Chen, CL Makino, NS Peachey, DA Baylor, and MI Simon. „Mechanisms of rhodopsin inactivation in vivo as revealed by a COOH-terminal truncation mutant.“ *Science* 267.5196 (Jan. 1995), pp. 374–7.

[139] BK Kobilka, C MacGregor, K Daniel, TS Kobilka, MG Caron, and RJ Lefkowitz. „Functional activity and regulation of human beta 2-adrenergic receptors expressed in Xenopus oocytes.“ *J. Biol. Chem.* 262.32 (Nov. 1987), pp. 15796–802.

[140] RL Dunbrack and FE Cohen. „Bayesian statistical analysis of protein side-chain rotamer preferences.“ *Protein Sci* 6.8 (Aug. 1997), pp. 1661–81. DOI: 10.1002/pro.5560060807.

[141] RC Rubenstein, SK Wong, and EM Ross. „The hydrophobic tryptic core of the beta-adrenergic receptor retains Gs regulatory activity in response to agonists and thiols.“ *J Biol Chem* 262.34 (Dec. 1987), pp. 16655–16662.

[142] PW Hildebrand, A Goede, RA Bauer, B Gruening, J Ismer, E Michalsky, and R Preissner. „SuperLooper–a prediction server for the modeling of loops in globular and membrane proteins.“ *Nucleic Acids Res.* 37.Web Server issue (2009), W571–4. DOI: 10.1093/nar/gkp338.

[143] C Kleuss and E Krause. „Galpha(s) is palmitoylated at the N-terminal glycine.“ *EMBO J.* 22.4 (Feb. 2003), pp. 826–32. DOI: 10.1093/emboj/cdg095.

[144] PB Wedegaertner, PT Wilson, and HR Bourne. „Lipid modifications of trimeric G proteins.“ *J. Biol. Chem.* 270.2 (Jan. 1995), pp. 503–6.

[145] B Hess, C Kutzner, D van der Spoel, and E Lindahl. „GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation.“ *J. Chem. Theory Comput.* 4.3 (Mar. 2008), pp. 435–447. DOI: 10.1021/ct700301q.

[146] MG Wolf, M Hoefling, C Aponte-Santamaría, H Grubmüller, and G Groenhof. „g_membed: Efficient insertion of a membrane protein into an equilibrated lipid bilayer with minimal perturbation.“ *J. Comput. Chem.* 31.11 (Aug. 2010), pp. 2169–74. DOI: 10.1002/jcc.21507.

[147] O Berger and O Edholm. „Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature.“ *Biophys. J.* 72.May 1997 (1997), pp. 2002–2013.

[148] HJC Berendsen, JR Grigera, and TP Straatsma. „The Missing Term in Effective Pair Potentials.“ *J. Phys. Chem.* 91.24 (1987), pp. 6269–6271.

[149] AW Schüttelkopf and DMF van Aalten. „PRODRG: a tool for high-throughput crystallography of protein-ligand complexes.“ *Acta Crystallogr. D. Biol. Crystallogr.* 60.Pt 8 (Aug. 2004), pp. 1355–63. DOI: 10.1107/S0907444904011679.

[150] C Kandt, J Schlitter, and K Gerwert. „Dynamics of water molecules in the bacteriorhodopsin trimer in explicit lipid/water environment.“ *Biophys. J.* 86.2 (2004), pp. 705–717.

[151] B Mertz, M Lu, MF Brown, and SE Feller. „Steric and electronic influences on the torsional energy landscape of retinal.“ *Biophys. J.* 101.3 (Aug. 2011), pp. L17–9. DOI: 10.1016/j.bpj.2011.06.020.

[152] G Bussi, D Donadio, and M Parrinello. „Canonical sampling through velocity rescaling.“ *J Chem Phys* 126.1 (2007), p. 14101.

[153] T Darden, D York, and L Pedersen. „Particle mesh Ewald: An N log(N) method for Ewald sums in large systems.“ en. *J. Chem. Phys.* 98.12 (June 1993), p. 10089. DOI: 10.1063/1.464397.

[154] JS Hub, BL de Groot, and D van der Spoel. „g_wham - A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates.“ *J. Chem. Theory Comput* 6.12 (Dec. 2010), pp. 3713–3720. DOI: 10.1021/ct100494z.

[155] W Kabsch and C Sander. „Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features.“ *Biopolymers* 22.12 (1983), pp. 2577–2637.

[156] W Kabsch. „A solution for the best rotation to relate two sets of vectors.“ *Acta Crystallogr. Sect. A* 32.5 (Sept. 1976), pp. 922–923. DOI: 10.1107/S0567739476001873.

[157] EA Coutsias, C Seok, and KA Dill. „Using quaternions to calculate RMSD.“ *J Comput Chem* 25.15 (2004), pp. 1849–1857.

[158] X Daura, K Gademann, B Jaun, WFV Gunsteren, AE Mark, D Seebach, and WF van Gunsteren. „Peptide Folding: When Simulation Meets Experiment.“ *Angew. Chem. Int. Ed* 38 (1999), pp. 236–240.

[159] E Espinosa, E Molins, and C Lecomte. „Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities.“ *Chem. Phys. Lett.* 285.3-4 (1998), pp. 170–173.

[160] JP Gallivan and Da Dougherty. „Cation-pi interactions in structural biology.“ *Proc Natl Acad Sci U S A* 96.17 (Aug. 1999), pp. 9459–64.

[161] ML Verdonk, JC Cole, MJ Hartshorn, CW Murray, and RD Taylor. „Improved protein-ligand docking using GOLD.“ *Proteins* 52.4 (Sept. 2003), pp. 609–23.

[162] C Altenbach, K Yang, DL Farrens, ZT Farahbakhsh, HG Khorana, and WL Hubbell. „Structural features and light-dependent changes in the cytoplasmic interhelical E-F loop region of rhodopsin: a site-directed spin-labeling study.“ *Biochemistry* 35.38 (Sept. 1996), pp. 12470–8. DOI: 10.1021/bi960849l.

[163] RU Malik, M Ritt, BT DeVree, RR Neubig, RK Sunahara, and S Sivaramakrishnan. „Detection of G protein-selective G protein-coupled receptor (GPCR) conformations in live cells.“ *J. Biol. Chem.* 288.24 (July 2013), pp. 17167–78. DOI: 10.1074/jbc.M113.464065.

[164] TH Kim, KY Chung, A Manglik, AL Hansen, RO Dror, TJ Mildorf, DE Shaw, BK Kobilka, and RS Prosser. „The Role of Ligands on the Equilibria Between Functional States of a.“ *J. Am. Chem. Soc.* 135.25 (June 2013), pp. 9465–74. DOI: 10.1021/ja404305k.

[165] RA Cerione, C Staniszewski, JL Benovic, RJ Lefkowitz, MG Caron, P Gierschik, R Somers, AM Spiegel, J Codina, and L Birnbaumer. „Specificity of the functional interactions of the beta-adrenergic receptor and rhodopsin with guanine nucleotide regulatory proteins reconstituted in phospholipid vesicles.“ *J. Biol. Chem.* 260.3 (Feb. 1985), pp. 1493–500.

[166] AS Rose, U Zachariae, H Grubmüller, KP Hofmann, P Scheerer, and PW Hildebrand. „The activating α5 helix switch is driven by the receptor G protein interface (working title, unpublished manuscript).“

[167] X Deupi, P Edwards, A Singhal, B Nickle, D Oprian, G Schertler, and J Standfuss. „Stabilized G protein binding site in the structure of constitutively active metarhodopsin-II.“ *Proc. Natl. Acad. Sci. U. S. A.* 109.1 (Jan. 2012), pp. 119–24.

[168] M Elgeti. „Spectroscopic Investigation of Visual Rhodopsin.“ Ph D thesis. Charité Berlin, 2012.

[169] M Mahalingam, K Martínez-Mayorga, MF Brown, and R Vogel. „Two protonation switches control rhodopsin activation in membranes.“ *Proc. Natl. Acad. Sci. U. S. A.* 105.46 (2008), pp. 17795–17800.

[170] B Knierim, KP Hofmann, OP Ernst, and WL Hubbell. „Sequence of late molecular events in the activation of rhodopsin.“ *Proc. Natl. Acad. Sci. U. S. A.* 104.51 (2007), pp. 20290–20295.

[171] S Nishimura, H Kandori, and A Maeda. „Interaction between photoactivated rhodopsin and the C-terminal peptide of transducin alpha-subunit studied by FTIR spectroscopy.“ *Biochemistry* 37.45 (Nov. 1998), pp. 15816–24. DOI: 10.1021/bi981451n.

[172] F Bartl, E Ritter, and KP Hofmann. „FTIR spectroscopy of complexes formed between metarhodopsin II and C-terminal peptides from the G-protein alpha- and gamma-subunits.“ *FEBS Lett.* 473.2 (2000), pp. 259–264.

[173] R Vogel, S Martell, M Mahalingam, M Engelhard, and F Siebert. „Interaction of a G protein-coupled receptor with a G protein-derived peptide induces structural changes in both peptide and receptor: a Fourier-transform infrared study using isotopically labeled peptides.“ *J. Mol. Biol.* 366.5 (2007), pp. 1580–1588.

[174] A Ansari, J Berendzen, SF Bownet, H Frauenfeldert, IET Iben, TB Sauke, E Shyamsunder, and RD Young. „Protein states and proteinquakes.“ *Proc. Natl. Acad. Sci. U. S. A.* 82.August (1985), pp. 5000–5004.

[175] H Frauenfelder, S Sligar, and P Wolynes. „The energy landscapes and motions of proteins.“ *Science (80-.).* 254.5038 (Dec. 1991), pp. 1598–1603. DOI: 10.1126/science.1749933.

[176] H Frauenfelder, G Chen, J Berendzen, PW Fenimore, H Jansson, BH McMahon, IR Stroe, J Swenson, and RD Young. „A unified model of protein dynamics.“ *Proc. Natl. Acad. Sci. U. S. A.* 106.13 (Mar. 2009), pp. 5129–34. DOI: 10.1073/pnas.0900336106.

[177] MA Wall, DE Coleman, E Lee, JA Iñiguez-Lluhi, BA Posner, AG Gilman, and SR Sprang. „The structure of the G protein heterotrimer Giα1β1γ2.“ *Cell* 83.6 (Dec. 1995), pp. 1047–1058. DOI: 10.1016/0092-8674(95)90220-1.

[178] JH Park, T Morizumi, Y Li, JE Hong, EF Pai, KP Hofmann, HW Choe, and OP Ernst. „Opsin, a structural model for olfactory receptors?“ *Angew. Chem. Int. Ed. Engl.* 52.42 (Oct. 2013), pp. 11021–4. DOI: 10.1002/anie.201302374.

[179] A Singhal, MK Ostermaier, Sa Vishnivetskiy, V Panneels, KT Homan, JJG Tesmer, D Veprintsev, X Deupi, VV Gurevich, GFX Schertler, and J Standfuss. „Insights into congenital stationary night blindness based on the structure of G90D rhodopsin.“ *EMBO Rep.* 14.6 (June 2013), pp. 520–6. DOI: 10.1038/embor.2013.44.

[180] Y Ashok, RT Nanekar, and VP Jaakola. *Crystallogenesis of adenosine A(2A) receptor-T4 lysozyme fusion protein: a practical route for the structure.* 1st ed. Vol. 520. Elsevier Inc., Jan. 2013, pp. 175–98. DOI: 10.1016/B978-0-12-391861-1.00008-3.

[181] M Ubbink. „The courtship of proteins: understanding the encounter complex.“ *FEBS Lett.* 583.7 (Apr. 2009), pp. 1060–6. DOI: 10.1016/j.febslet.2009.02.046.

[182] B Ma, S Kumar, CJ Tsai, and R Nussinov. „Folding funnels and binding mechanisms.“ *Protein Eng. Des. Sel.* 12.9 (Sept. 1999), pp. 713–720. DOI: 10.1093/protein/12.9.713.

[183] Y Levy and JN Onuchic. „Water mediation in protein folding and molecular recognition.“ *Annu. Rev. Biophys. Biomol. Struct.* 35 (2006), pp. 389–415.

[184] TW Schwartz and TP Sakmar. „Structural biology: Snapshot of a signalling complex.“ *Nature* 477.7366 (2011), pp. 540–541.

[185] R Rahmeh, M Damian, M Cottet, H Orcel, C Mendre, T Durroux, KS Sharma, G Durand, B Pucci, E Trinquet, JM Zwier, X Deupi, P Bron, JL Banères, B Mouillac, and S Granier. „Structural insights into biased G protein-coupled receptor signaling revealed by fluorescence spectroscopy.“ *Proc. Natl. Acad. Sci. U. S. A.* 109.17 (Apr. 2012), pp. 6733–8. DOI: 10.1073/pnas.1201093109.

[186] Ma Soriano-Ursúa, JG Trujillo-Ferrara, J Correa-Basurto, and S Vilar. „Recent structural advances of β1 and β2 adrenoceptors yield keys for ligand recognition and drug design.“ *J. Med. Chem.* 56.21 (Nov. 2013), pp. 8207–23. DOI: 10.1021/jm400471z.

[187] O Kisselev, C Meyer, M Heck, OP Ernst, and KP Hofmann. „Signal transfer from rhodopsin to the G-protein: evidence for a two-site sequential fit mechanism.“ *Proc Natl Acad Sci U S A* 96.9 (1999), pp. 4898–4903.

[188] WM Oldham, N Van Eps, AM Preininger, WL Hubbell, and HE Hamm. „Mapping allosteric connections from the receptor to the nucleotide-binding pocket of heterotrimeric G proteins.“ *Proc. Natl. Acad. Sci. U. S. A.* 104.19 (2007), pp. 7927–7932.

[189] G Schreiber. „Kinetic studies of protein-protein interactions.“ *Curr. Opin. Struct. Biol.* 12.1 (Feb. 2002), pp. 41–7.

[190] Jy Suh, C Tang, and GM Clore. „Role of electrostatic interactions in transient encounter complexes in protein-protein association investigated by paramagnetic relaxation enhancement.“ *J. Am. Chem. Soc.* 129.43 (Oct. 2007), pp. 12954–5. DOI: 10.1021/ja0760978.

[191] OP Ernst, CK Meyer, EP Marin, P Henklein, WY Fu, TP Sakmar, and KP Hofmann. „Mutation of the Fourth Cytoplasmic Loop of Rhodopsin Affects Binding of Transducin and Peptides Derived from the Carboxyl-terminal Sequences of Transducin and Subunits.“ *J. Biol. Chem.* 275.3 (Jan. 2000), pp. 1937–1943. DOI: 10.1074/jbc.275.3.1937.

[192] VV Gurevich and JL Benovic. „Visual arrestin interaction with rhodopsin. Sequential multisite binding ensures strict selectivity toward light-activated phosphorylated rhodopsin.“ *J. Biol. Chem.* 268.16 (June 1993), pp. 11628–38.

[193] K Schröder, A Pulvermüller, and KP Hofmann. „Arrestin and its splice variant Arr1-370A (p44). Mechanism and biological role of their interaction with rhodopsin.“ *J. Biol. Chem.* 277.46 (Nov. 2002), pp. 43987–96. DOI: 10.1074/jbc.M206211200.

[194] D Fotiadis, Y Liang, S Filipek, DA Saperstein, A Engel, and K Palczewski. „Atomic-force microscopy: Rhodopsin dimers in native disc membranes.“ *Nature* 421.6919 (Jan. 2003), pp. 127–8. DOI: 10.1038/421127a.

[195] Y Liang, D Fotiadis, S Filipek, Da Saperstein, K Palczewski, and A Engel. „Organization of the G protein-coupled receptors rhodopsin and opsin in native membranes.“ *J. Biol. Chem.* 278.24 (June 2003), pp. 21655–62. DOI: 10.1074/jbc.M302536200.

[196] D Dell'Orco. „A physiological role for the supramolecular organization of rhodopsin and transducin in rod photoreceptors.“ *FEBS Lett.* 587.13 (June 2013), pp. 2060–6. DOI: 10.1016/j.febslet.2013.05.017.

[197] J Schöneberg, M Heck, KP Hofmann, and F Noé. „Explicit spatiotemporal simulation of receptor-G protein coupling in rod cell disk membranes.“ *Biophys. J.* 107.5 (Sept. 2014), pp. 1042–53. DOI: 10.1016/j.bpj.2014.05.050.

[198] JA Hirsch, C Schubert, VV Gurevich, and PB Sigler. „A Model for Arrestin's Regulation: The 2.8 ÅCrystal Structure of Visual Arrestin.“ *Cell* 97.2 (Apr. 1999), pp. 257–269. DOI: 10.1016/S0092-8674(00)80735-7.

[199] AK Shukla, A Manglik, AC Kruse, K Xiao, RI Reis, WC Tseng, DP Staus, D Hilger, S Uysal, LY Huang, M Paduch, P Tripathi-Shukla, A Koide, S Koide, WI Weis, Aa Kossiakoff, BK Kobilka, and RJ Lefkowitz. „Structure of active β-arrestin-1 bound to a G-protein-coupled receptor phosphopeptide.“ *Nature* 497.7447 (May 2013), pp. 137–41. DOI: 10.1038/nature12120.

[200] J Schilder and M Ubbink. „Formation of transient protein complexes.“ *Curr. Opin. Struct. Biol.* 23.6 (Dec. 2013), pp. 911–8. DOI: 10.1016/j.sbi.2013.07.009.

[201] MJM Niesen, S Bhattacharya, and N Vaidehi. „The Role of Conformational Ensembles in Ligand Recognition in G-Protein Coupled Receptors.“ *J. Am. Chem. Soc.* 133.33 (2011), pp. 13197–13204.

[202] R Aurora and GD Rose. „Helix capping.“ *Protein Sci* 7.1 (Jan. 1998), pp. 21–38.

[203] IK McDonald and JM Thornton. „Satisfying hydrogen bonding potential in proteins.“ *J Mol Biol* 238.5 (1994), pp. 777–793.

[204] AC Wallace, RA Laskowski, and JM Thornton. „LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions.“ *Protein Eng* 8.2 (1995), pp. 127–134.

[205] J Li, PC Edwards, M Burghammer, C Villa, and GFX Schertler. „Structure of bovine rhodopsin in a trigonal crystal form.“ *J. Mol. Biol.* 343.5 (2004), pp. 1409–1438.

[206] RRR Neubig, M Spedding, T Kenakin, and A Christopoulos. „Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification. XXXVIII. Update on terms and symbols in quantitative pharmacology.“ *Pharmacol. ...* 55.4 (Dec. 2003), pp. 597–606. DOI: 10.1124/pr.55.4.4.and.

1AZT	PDB entry of $G\alpha_s$ complexed with GTPγS; X-ray crystal structure at 2.3 Å resolution. [65]
1GP2	PDB entry of G protein heterotrimer $G\alpha_i\beta\gamma$ with GDP bound; X-ray crystal structure at 2.3 Å resolution. [177]
1GZM	PDB entry of dark state/inactive rhodopsin; trigonal crystal form; X-ray crystal structure at 2.65 Å resolution. [205]
1U19	PDB entry of dark state/inactive rhodopsin; X-ray crystal structure at 2.2 Å resolution. [39]
2RH1	PDB entry of inactive β_2AR; X-ray crystal structure at 2.4 Å resolution. [40]
3CAP	PDB entry of active opsin; X-ray crystal structure at 2.9 Å resolution. [30]
3DQB	PDB entry of active opsin; with bound high affinity $G\alpha_t$CT; X-ray crystal structure at 3.2 Å resolution. [31]
3P0G	PDB entry of nanobody stabilized active β_2AR; X-ray crystal structure at 3.5 Å resolution. [60]
3PQR	PDB entry of metarhodopsin II; with bound double high affinity $G\alpha_t$CT; X-ray crystal structure at 2.85 Å resolution. [66]
3PXO	PDB entry of metarhodopsin II, X-ray crystal structure at 3.0 Å resolution. [66]
3SN6	PDB entry of β_2AR*•$G\alpha_s\beta\gamma$ complex; X-ray crystal structure at 3.2 Å resolution. [32]
ABMD	adiabatic biased MD
agonist	A ligand that increases receptor activity [206].
CaPTURE	A software program that detects cation-π interactions in protein structures by estimating the electrostatic and the van der Waals interaction energy [160].
DEER	double electron-electron resonance
DMPC	*dimyristoylphosphatidylcholine*
DOWSER	A software tool that assesses if an internal water molecule makes a significant energetic contribution to the structural stability by interacting with their surrounding atoms [133].
DSSP	The standard algorithm for defining the secondary structure of proteins that are available structures at atomic resolution. Based on electrostatic interactions and the geometry of the protein backbone, the algorithm assigns eight types of secondary structure, including α-helices and β-sheets [155].
ECL	extra-cellular loop
EM	electron microscopy

EPR	electron paramagnetic resonance
FMM	fast multipole method
FTIR	Fourier transform infrared spectroscopy
GPCR	G protein-coupled receptor
GPU	graphics processing unit
GROMACS	A popular software package to perform and analyze molecular dynamics simulations [145].
HDX	hydrogen-deuterium exchange
ICL	intra-cellular loop
inverse agonist	A ligand that reduces receptor activity, e. g. by stabilizing the inactive receptor state [206].
JSON	JavaScript object notation. A lightweight data-interchange format (`http://www.json.org/`).
ligand	A molecule that alters the activity of a receptor [206].
MD	Molecular dynamics (Section 2.1).
MSM	multilevel summation method
NMR	nuclear magnetic resonance
PME	particle-mesh Ewald
PMF	potential of mean force
Python	A high-level, multi-purpose programming language (`http://www.python.org`).
RMSD	Root mean square deviation (Section 2.4.2).
RMSF	Root mean square fluctuation (Section 2.4.2).
SAS	solvent accessible surface
SDSL	site-directed spin labeling
TM	trans-membrane helix
US	Umbrella sampling (Section 2.3.3).

Selbständigkeitserklärung

Hiermit erkläre ich, die Dissertation selbstständig und nur unter Verwendung der angegebenen Hilfen und Hilfsmittel angefertigt zu haben.

Berlin, den 18.12.2014

_____

Alexander S. Rose